❧ NCBI   ≡Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM

Search [Nucleotide ▼] for [   ] [Go] [Clear]

Limits   Index   History   Clipboard

[Display] [Default View ▼] as [HTML ▼] [Save] [Add to Clipboard]

☐ 1: <u>AL034446</u> **Streptomyces coelicolor cosmid 1A9** PubMed, Protein, Related Sequences, Taxonomy

```
LOCUS       SC1A9          29400 bp    DNA          BCT        11-DEC-1998
DEFINITION  Streptomyces coelicolor cosmid 1A9.
ACCESSION   AL034446
VERSION     AL034446.1  GI:4007685
KEYWORDS    ADA-like regulatory protein; araC family; gntR; helix-turn-helix;
            His rich; integral membrane; luxR; luxR family;
            methylated-DNA-protein-cysteine methyl transferase; nucleotide
            binding; ogt; oxidoreductase; poxB; pyruvate oxidase; response
            regulator; sensor kinase; sorbitol oxid ase; thiamine
            pyrophosphate; TPP; transferase; transmembrane protein; two
            component system.
SOURCE      Streptomyces coelicolor A3(2).
  ORGANISM  Streptomyces coelicolor A3(2)
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 29400)
  AUTHORS   Saunders,D.C. and Harris,D.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29400)
  AUTHORS   Bentley,S.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1998) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
REFERENCE   3  (bases 1 to 29400)
  AUTHORS   Redenbach,M., Kieser,H.M., Denapaite,D., Eichner,A., Cullum,J.,
            Kinashi,H. and Hopwood,D.A.
  TITLE     A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
  JOURNAL   Mol. Microbiol. 21 (1), 77-96 (1996)
  MEDLINE   97000351
COMMENT     Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC.
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
            numbered using the following system eg SC7B7.01c. SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous. The length
            in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon. Gene
            prediction is based on positional base preference in codons using a
            specially developed Hidden Markov Model (Krogh et al., Nucleic
```

Acids Research, 22(22):4768-4778(1994)) and the FramePlot program
of Bibb et al., Gene 30:157-66(1984) as implemented at
http://www.nih.go.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gtg, ttg or (att)) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions. Cosmid 1A9 lies
between 9B2 and 6C5 in the Ase-I-B genomic restriction fragment.

```
FEATURES             Location/Qualifiers
     source          1..29400
                     /organism="Streptomyces coelicolor A3(2)"
                     /strain="A3(2)"
                     /db_xref="taxon:100226"
                     /clone="cosmid 1A9"
     gene            complement(1..651)
                     /gene="SC1A9.01c"
     CDS             complement(<1..651)
                     /gene="SC1A9.01c"
                     /note="SC1A9.01c, incomplete CDS, possible transferase,
                     partial CDS, len: 219 aa, similar to TR:O53185
                     (EMBL:AL021246) a proposed transferase from Mycobacterium
                     tuberculosis (302 aa), fasta scores: opt:  764, z-score:
                     1097.3, E(): 0, (62.1% identity in 203 aa overlap (302
                     aa). Also weakly similar to several methyltransferases eg.
                     METH_ECOLI
                     (EMBL:X16584)5-methyltetrahydrofolate-homocysteine
                     methyltransferase (1226 aa), fasta scores; opt:160,
                     z-score: 234.8, E(): 8.4e-06, (30.0% identity in 220 aa
                     overlap)."
                     /codon_start=1
                     /transl_table=11
                     /product="putative transferase"
                     /protein_id="CAA22371.1"
                     /db_xref="GI:4007686"
                     /translation="MTSDFADALASGPLVLDGGLSNQLEAAGHDLGDALWSARLLAED
                     PEAITRAHLAYFEAGAEVAITSSYQATFEGFARRGIGRERAAELLALSVASAREAARR
                     ARTARPERALWVAASAGPYGAMLADGSEYRGRYGLGRGALERFHRPRLEVLAAARPDV
                     LALETVPDTDEAAALLRAVRGLDVPAWLSYTVAGDRTRAGQPLDEAFALAADVDEVI"
     gene            731..1654
                     /gene="SC1A9.02"
     CDS             731..1654
                     /gene="SC1A9.02"
                     /note="SC1A9.02, possible transmembrane protein, len:
                     307aa; Contains several possible membrane spanning
                     domains."
                     /codon_start=1
                     /transl_table=11
                     /product="putative transmembrane protein"
                     /protein_id="CAA22372.1"
                     /db_xref="GI:4007687"
                     /translation="MTISGRRIRSVRCSPRHSSHGPDEGVPVRFVWQFLAVLVAYAIG
                     GIAVQAVKDNDWLTLVVGLTSVALVVFVYAWVVRRTERREALDVALDGAAAKAGWGTL
                     IGFGLFGAVITNLFASGYYEVDGLGSVQGAIGLVGFMAAAAATEEVVFRGVLFRIIEE
                     HIGTYLALGLTGLVFGLMHLLNEDATLWGALAIAIEAGFMLAAAYAATRNLWLTIGVH
                     FGWNFAAGGVFSTVVSGNGDSEGLLDATMSGPKLLTGGDFGPEGSVYSVGFGVLLTLV
                     FLWLAHRRGNIVAFGSRRRAAGANSAATLPR"
     gene            1651..2817
                     /gene="SC1A9.03"
     CDS             1651..2817
```

```
                        /gene="SC1A9.03"
                        /note="SC1A9.03, probable two component sensor kinase,
                        len: 388aa; similar to a family of sensor kinases egs.
                        TR:Q53893 (EMBL:U51332) AbsA1 from Streptomyces coelicolor
                        (571 aa), fasta scores; opt: 389, z-score: 278.9, E():
                        2.9e-08, (33.1% identity in 405 aa overlap) and UHPB_ECOLI
                        (EMBL:M17102) sensor kinase from Escherichia coli (500
                        aa), fasta scores; opt:221, z-score: 248.0, E(): 1.5e-06,
                        (25.9% identity in 398 aa overlap). Contains several
                        possible membrane spanning domains."
                        /codon_start=1
                        /transl_table=11
                        /product="putative sensor kinase"
                        /protein_id="CAA22373.1"
                        /db_xref="GI:4007688"
                        /translation="MIDRRRVLELWRRLDVTVRDLPLGVLLLVASLLPSLRGQGTEIG
                        GLPTRPADALAGVAAVLQSIPLAVRRRWTLLCLTLVSLGFALDQLRAYHLFAGAALPI
                        VLINAGSHQEKYRRATQVTATLGYVAMAVGLNARGGDETLVEYVTFYLVLALAWGIGA
                        WMRSARAAEAERRSRVAEDARNAERTRIARELHDVVTHHVTAMVVQSEAARYLTAAPE
                        RLDESLAAVSDTGRRAITDLRHLLDLLNPDHGTAEPRTPPVGRVLTLVEQTRRAGQPV
                        EFTEEGTPAAATGSSDLVAYRVVQEALTNALKYDHGGRTSVLVRHGEREITVEVGTDG
                        SGSGAASPGGSGRGLAGLRERVDVLGGEFSTDRPADGGFVVRARIPGGSGGSTA"
gene                    2814..3473
                        /gene="SC1A9.04"
CDS                     2814..3473
                        /gene="SC1A9.04"
                        /note="SC1A9.04, probable luxR family response regulator,
                        len: 219 aa; similar to a family of regulators egs.
                        TR:O69816 (EMBL:AL023496) probable two-component regulator
                        from Streptomyces coelicolor (224 aa) fasta scores;
                        opt:714, z-score:1103.5, E():0, (54.8% identity in 219 aa
                        overlap) and NARL_ECOLI (EMBL:X13360) nitrate /nitrite
                        response regulator from Escherichia coli (216 aa) fasta
                        scores; opt: 465, z-score: 656.5, E(): 2.7e-29, (37.6%
                        identity in 213 aa overlap). Contains Pfam match to entry
                        PF00072 response_reg, Response regulator receiver domain,
                        score 111.10, E-value 2.1e-29, Pfam match to entry PF00196
                        GerE, Bacterial regulatory proteins, luxR family, score
                        79.50, E-value 7e-20, PS00622 Bacterial regulatory
                        proteins, luxR family signature and an helix-turn-helix
                        motif from: 1 to: 219, Score 983 (+2.53 SD)."
                        /codon_start=1
                        /transl_table=11
                        /product="putative response regulator"
                        /protein_id="CAA22374.1"
                        /db_xref="GI:4007689"
                        /translation="MSAPIRVVICDDQALIRTGLATIVDAQPDLEVVGECGDGQTGVD
                        LARELRPDVVVMDIRMPVLDGLEATRLLAGAGVAHPVKVLVVTTFNLDEYVYEALRAG
                        ASGFLLKDAPPDRLLHGIRTVAMGAALLDPDVTRRLVGRYAARIRPAEGTARDIPLTP
                        RETEVLRLIADGLSNSEIAAALVISPETVKTFVSRILTKLDLRDRVQAVVFAYRHGLV
                        T"
misc feature           2826..3173
                        /gene="SC1A9.04"
                        /note="Pfam match to entry PF00072 response_reg, Response
                        regulator receiver domain, score 111.10, E-value 2.1e-29"
misc feature           3276..3470
                        /gene="SC1A9.04"
                        /note="Pfam match to entry PF00196 GerE, Bacterial
                        regulatory proteins, luxR family, score 79.50, E-value
                        7e-20"
misc feature           3327..3410
                        /gene="SC1A9.04"
                        /note="PS00622 Bacterial regulatory proteins, luxR family
                        signature."
gene                    complement(3492..4094)
```

```
     CDS         /gene="SC1A9.05c"
                 complement(3492..4094)
                 /gene="SC1A9.05c"
                 /note="SC1A9.05c, possible transmembrane protein, len: 200
                 aa; contains possible membrane spanning hydrophobic
                 regions"
                 /codon_start=1
                 /transl_table=11
                 /product="putative transmembrane protein"
                 /protein_id="CAA22375.1"
                 /db_xref="GI:4007690"
                 /translation="MFRANAGIRSAPYVHVFPMRGDMASTERSDDRPRLRRRDSLWGI
                 GVMALLAGVVVRLVLNGTSAWLSALLGAVPAAVWIVWWVRRRRVRDARAVGAEPDDVP
                 AMERQILKGGPAPRDPERRRAMAAFVDSRQERLRRNRWWAFPMLAVIFFGTSALWYLS
                 GSVGAGSLMLGLGVVFLGWLAWYNLRIDRRLSHMRGRLRG"
     gene        4140..4631
                 /gene="SC1A9.06"
     CDS         4140..4631
                 /gene="SC1A9.06"
                 /note="SC1A9.06, unknown, len: 163 aa"
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein SC1A9.06"
                 /protein_id="CAA22376.1"
                 /db_xref="GI:4007691"
                 /translation="MRDMNDTRTPTSTVTSTRTPADAVTGMVDHVLHLAAGWTRWDGT
                 PAHVDGRVYTPHKAIRRVADHLVDHLAELEARLAGEETQPDHWHASLVTTEADRAAFT
                 AEDLDEARSRLTRLARIWANRLDALTDEQLDHSPGEGWSFRELAAHLAESGYYADAVG
                 DLS"
     gene        4628..5383
                 /gene="SC1A9.07"
     CDS         4628..5383
                 /gene="SC1A9.07"
                 /note="SC1A9.07, unknown, len: 251 aa; similar to
                 TR:O33986 (EMBL:U82823) hypothetical protein from
                 Saccharopolyspora erythraea (266 aa), fasta scores; opt:
                 740, z-score: 869.0, E(): 0, (53.4% identity in 251 aa
                 overlap)."
                 /codon_start=1
                 /transl_table=11
                 /product="hypothetical protein SC1A9.07"
                 /protein_id="CAA22377.1"
                 /db_xref="GI:4007692"
                 /translation="MTATAFAALHRAGEPLLLPCAWDHASAFALAGQGFRAVGTTSLG
                 VAAAAGLPDGASATRDETLRLALVLGSAPFPLSVDAEDGFSDDPDEVGEFARQLAAVG
                 AVGINLEDGLGPVGRHAAKIAAVRSAAPGLFVNARTDTYWSGDGDVTETLRRLEAYRE
                 AGADGVFVPGLTDPARIGSLAARFDVPLNVLYTPAGPGLAHLADLGVRRVSLGSLLYR
                 RALGAALRAAADVRAGRDPGGPTPTYDEVRAPG"
     gene        complement(5386..5955)
                 /gene="SC1A9.08c"
     CDS         complement(5386..5955)
                 /gene="SC1A9.08c"
                 /note="SC1A9.08c, possible transcriptional regulator, len:
                 189 aa; similar to TR:O34892 (EMBL:AF027868) proposed
                 transcriptional regulator from Bacillus subtilis (191 aa)
                 fasta scores; opt: 302, z-score: 488.2, E(): 6.6e-20,
                 (31.8% identity in 173 aa overlap). Contains an
                 helix-turn-helix motif from: 1 to: 189, Score 1074 (+2.84
                 SD)."
                 /codon_start=1
                 /transl_table=11
                 /product="putative transcriptional regulator"
                 /protein_id="CAA22378.1"
                 /db_xref="GI:4007693"
                 /translation="MPRVGLTTDRVVAAAADLADETGFESVTVSALARHFGVKDASLY
```

```
                               THVRNLQDLRVRVALLAGGELIEEIAGAVAGRAGKEALAAFAGAYRAYALRHPGRYAA
                               TQIRVDQSLVADSAALRRTAEITYGMLRSYGLTEPDLTDAVRLLRSTFHGYCALESSG
                               AFGAPRDVRASWDKAVDALHVALENWPRA"
              gene             6107..7825
                               /gene="SC1A9.09"
              CDS              6107..7825
                               /gene="SC1A9.09"
                               /note="SC1A9.09, unknown, len: 572 aa; Contains His rich
                               extreme N-terminal."
                               /codon_start=1
                               /transl_table=11
                               /product="hypothetical protein SC1A9.09"
                               /protein_id="CAA22379.1"
                               /db_xref="GI:4007694"
                               /translation="MGHGHGHPHGHHHHGHTHDHEHATEQVLPAAFDTSVPDEALSPA
                               QQSRRGLLRRAGLLGAGLAAGTVLAPAATATAAPARAASNGRRGKGFLWLAGDHHIHT
                               QYSSDGKYRVVDQVRQGARHGMDWLVITDHGSNTHAKIGVEKVNPDIREARAAHEDTL
                               VFQGLEWNIPAAEHGTVFVHPGKHEVSVLKQFETDYDGSVKGAGDSTPANEALAIAGL
                               SFLADQVKRRKVKDALMLANHPARKGIDSPHEIRAWRDATSRGHQIAVGFEGAPGHQA
                               GGLPEPLGPGGARGIYDGSPSANSFAGYPLESYRTWGGFDWMTATVGGLWDSLIAEGR
                               PWWITANSDSHQVYADTGARGGGDFNANGRYDDPVYAGQIDITQNDFWPGQYSRTHVG
                               SDGFSYAAVMDGIRAGRVWVDHGQLVSGLDVRVSGGGRWATLGGALHVRRGTRVTLSI
                               DVALAGGPNWAGFVPKLARVDVIQGDVTGPAADKDTFTAPTARVVKSYEVDKETGTVR
                               LTYDLGRVDRPVYLRTRGTDGNRSAVGSLGAKVDPAGPAIDVVGDADPWRDLWFYSNP
                               VWVLPS"
              gene             7822..8346
                               /gene="SC1A9.10"
              CDS              7822..8346
                               /gene="SC1A9.10"
                               /note="SC1A9.10, unknown, len: 174 aa"
                               /codon_start=1
                               /transl_table=11
                               /product="hypothetical protein SC1A9.10"
                               /protein_id="CAA22380.1"
                               /db_xref="GI:4007695"
                               /translation="MTPPAAPYVLGIDTDATTLREADHLLQALAAELDLPEGVFGCTH
                               LVRDGRPRVALSLAAEAEPVLRTARDRLTARGHEVRDGTWDESGRAVLFPGAAALTGT
                               LTLAELLARSAVDRVTVLGTPDEPSPDTRLVTRNHVRPHWQDGRLVLAAMPAVGGTLV
                               PFEDPDPTPCCADH"
              gene             complement(8377..9633)
                               /gene="SC1A9.11c"
              CDS              complement(8377..9633)
                               /gene="SC1A9.11c"
                               /note="SC1A9.11c, oxidoreductase, len: 418 aa, similar to
                               many eg. TR:P97011 (EMBL:AB000519) proposed sorbitol
                               oxidase from Streptomyces sp. (420 aa) fasta scores; opt:
                               1544, z-score: 1630.2, E():0, (60.0% identity in 413 aa
                               overlap) and TR:O50531 (EMBL:AL009204) FAD-dependent
                               oxidoreductase from Streptomyces coelicolor (445 aa) fasta
                               scores; opt: 374, z-score: 499.4, E(): 1.6e-20, (31.4%
                               identity in 439 aa overlap)."
                               /codon_start=1
                               /transl_table=11
                               /product="putative oxidoreductase"
                               /protein_id="CAA22381.1"
                               /db_xref="GI:4007696"
                               /translation="MSDITVTNWAGNITYTAKELLRPHSLDALRALVADSARVRVLGS
                               GHSFNEIAEPGDGGVLLSLAGLPSVVDVDTAARTVRVGGGVRYAELARVVHARGLALP
                               NMASLPHISVAGSVATGTHGSGVGNGSLASVVREVELVTADGSTVVIARGDERFGGAV
                               TSLGALGVVTSLTLDLEPAYEMEQHVFTELPLAGLDPATFETVMAAAYSVSLFTDWRA
                               PGFRQVWLKRRTDRPLDGFPYAAPAAEKMHPVPGMPAVNCTEQFGVPGPWHERLPHFR
                               AEFTPSSGAELQSEYLMPREHALAALHAMDAIRETLAPVLQTCEIRTVAADAQWLSPA
                               YGRDTVAAHFTWVEDTAAVLPVVRRLEEALVPFAARPHWGKVFTVPAGELRALYPRLA
                               DFGALAGALDPAGKFTNAFVRGVLAG"
              gene             9740..11068
```

```
                    /gene="SC1A9.12"
        CDS         9740..11068
                    /gene="SC1A9.12"
                    /note="SC1A9.12, unknown, len: 442 aa"
                    /codon_start=1
                    /transl_table=11
                    /product="hypothetical protein SC1A9.12"
                    /protein_id="CAA22382.1"
                    /db_xref="GI:4007697"
                    /translation="MDSRTALVEDLMERFPHVPREAVFKEDLLRGGVAFDPSALSDTT
                    NEAAGEVKPKSYFIFSFDHGTLPELGEAALRRPPEEIILTGGPYDLRRTVVSVRVNPA
                    SPYRVAANEDGVLGLYLDGKRIADVGVPPMPEYYRHKLSNGKSVMEVAPTIQWGYLIY
                    LTAFRVCQYFGAKEECQYCDINHNWRQHKAAGRPYTGVKDVDEVLEALEIIDKYDTAK
                    ISTAYTLTGGAITSKVQGLDEADFYGRYAKAIEEHFPGRWIGKVVAQALPKPDVQRFK
                    DYGVQIYHPNFEVWDEYLFKMYCPGKERYVGRDEWHKRILDSTEVFGARNVIPNFVAG
                    VEMAEPFGFKTVDEAIESTTEGLRFFMSHGITPRFTTWCPEPTTPLGKTNPDGAPLEY
                    HIRLLQAYRQTMEDYGLSSPPGYGPPGAGNAVFSVSSFMDSLPEDAPVEV"
        gene        11316..12416
                    /gene="SC1A9.13"
        CDS         11316..12416
                    /gene="SC1A9.13"
                    /note="SC1A9.13, possible nucleotide binding protein, len:
                    366 aa; similar to several hypotheticals eg. YJEQ_HAEIN
                    (EMBL:U32844) hypothetical protein HI1714 from Haemophilus
                    influenzae (346 aa) fasta scores; opt:472, z-score: 564.4,
                    E(): 3.7e-24, (31.0% identity in 313 aa overlap). Contains
                    PS00017 ATP /GTP-binding site motif A (P-loop)."
                    /codon_start=1
                    /transl_table=11
                    /product="putative integral membrane nucleotide binding
                    protein"
                    /protein_id="CAA22383.1"
                    /db_xref="GI:4007698"
                    /translation="MTSTSSHSNHSALSSYGWDDSWADAFAPYAAEGLLPGRVVRVDR
                    GQCDVVTADGVLRADTAFVTPHDPLRVVCTGDWVAVEPGGNPRYVRTYLPRRTAFVRS
                    TSSKRSEGQILAANVDHAVVAVSLAVELDLARIERFLALAWESGAQPLVVLTKADLVP
                    DPVTLAYLVQDVETAAPGVPVLPVSAEQGEGLDVLAAVVSGGTAVLLGQSGAGKSTLA
                    NALLGEAAMDVQAIRDVDGKGRHTTTTRNLLALPGGGVLIDTPGLRGVGLFDAGNGVD
                    QVFAEIAELAEECRFHDCAHESEPGCAVLAAIDSGALPERRLESYRKLMRENQRIVAK
                    TDARARAEIRKEYKRRGAIGKAAMEAKRGGLR"
        misc_feature 11937..11960
                    /gene="SC1A9.13"
                    /note="PS00017 ATP /GTP-binding site motif A (P-loop)"
        gene        12489..13922
                    /gene="SC1A9.14"
        CDS         12489..13922
                    /gene="SC1A9.14"
                    /note="SC1A9.14, possible ADA-like regulatory protein,
                    len: 477 aa; similar to ADA_MYCTU (EMBL:Z73902) putative
                    ADA regulatory protein from Mycobacterium tuberculosis
                    (496 aa) fasta scores; opt: 1135, z-score: 1376.3, E(): 0,
                    (51.6% identity in 486 aa overlap). Also note N-terminal
                    region similar to  N-terminal region of ADA_ECOLI
                    (EMBL:M10211) ADA regulatory protein from Escherichia coli
                    (354 aa) fasta scores; opt: 302, z-score: 255.8, E():
                    5.7e-07, (34.5% identity in 171 aa overlap) and C-terminal
                    region similar to 3MG2_ECOLI AlkA, 3-methyladenine DNA
                    glycosylase II involved in DNA repair (282 aa) fasta
                    scores; opt: 171, z-score: 232.9, E(): 1.1e-05, (30.0%
                    identity in 290 aa overlap). Contains Pfam match to entry
                    PF00165 HTH_2, Bacterial regulatory helix-turn-helix
                    proteins, araC family, score 72.90, E-value 6.7e-18 and
                    PS00041 Bacterial regulatory proteins, araC family
                    signature."
                    /codon_start=1
```

```
                        /transl_table=11
                        /product="putative ADA-like regulatory protein"
                        /protein_id="CAA22384.1"
                        /db_xref="GI:4007699"
                        /translation="MTPQTVQPAEHADAREDVRYEAVRSRDARFDGAFFFAVETTGIY
                        CRPSCPAVTPKRRNVRFFATAAAAQGSGFRACRRCRPDAVPGSADWNVRADVVGRAMR
                        LIGDGVVDREGVAGLAGRLGYSARQVQRQLTAEVGAGPVALARAQRAHTARVLLQTTV
                        LPVTEIAFASGFASVRQFNDTIRAVYAATPSELRAAAPARDRAARRTATPSAGVPLRL
                        AHRGPYQAGPVFDLLQREAVTGVEEVSGETGRRLYRRTLRLPYGTGIVAVQERPGRAG
                        TGSGGWLEARLHLTDLRDLTTSVQRLRRLFDLDADPYAVDERLGADPRLAPLVAARPG
                        LRSPGTADPAELAVRALVGRTEAERLVQRYGKALDAPCGTLTHLFPEPDVLAGAAPHG
                        TPGALAAALADGAVRLDPGADRDDAERALLAVPGLDARTVAVVRTRALGDPDVAPPGA
                        AVPDTWRPWRSYALNHLRAAGEWENDR"
     misc_feature       12819..13079
                        /gene="SC1A9.14"
                        /note="Pfam match to entry PF00165 HTH_2, Bacterial
                        regulatory helix-turn-helix proteins, araC family, score
                        72.90, E-value 6.7e-18"
     misc_feature       12930..13058
                        /gene="SC1A9.14"
                        /note="PS00041 Bacterial regulatory proteins, araC family
                        signature."
     gene               13919..14479
                        /gene="SC1A9.15"
     CDS                13919..14479
                        /gene="SC1A9.15"
                        /note="SC1A9.15, ogt, methylated-DNA-protein-cysteine
                        methyltransferase len: 186 aa; similar to many eg.
                        DAT1_BACSU (EMBL:X15659) methylated-DNA-protein-cysteine
                        methyltransferase from Bacillus subtilis (165 aa) fasta
                        scores; opt:357, z-score: 478.2, E(): 2.3e-19, (41.4%
                        identity in 157 aa overlap). Contains Pfam match to entry
                        PF01035 Methyltrans, 6-O-methylguanine DNA
                        methyltransferase, score 170.70, E-value 2.5e-47."
                        /codon_start=1
                        /transl_table=11
                        /product="putative methylated-DNA-protein-cysteine
                        methyltransferase"
                        /protein_id="CAA22385.1"
                        /db_xref="GI:4007700"
                        /translation="MTTTTPTTTTTSIPAETYWHEVDSPVGPLLLTAGSDGALTSLSV
                        PGQKGGRSVRDGWRHDAGPFRVAEEQLGAYFAGELTEFSLPLRAQGTAFRERVWAALD
                        DVPYGATTTYGEIAARIGASRPAVRAVGGAIGANPLLILRPCHRVIGADGSLTGYAGG
                        LERKTRLLSLEGAPLSRPVPLPATPR"
     misc_feature       14111..14446
                        /gene="SC1A9.15"
                        /note="Pfam match to entry PF01035 Methyltrans,
                        6-O-methylguanine DNA methyltransferase, score 170.70,
                        E-value 2.5e-47"
     gene               complement(14442..14924)
                        /gene="SC1A9.16c"
     CDS                complement(14442..14924)
                        /gene="SC1A9.16c"
                        /note="SC1A9.16c, small hydrophobic protein, len: 160 aa"
                        /codon_start=1
                        /transl_table=11
                        /product="small hydrophobic protein"
                        /protein_id="CAA22386.1"
                        /db_xref="GI:4007701"
                        /translation="MGAWDLLLAGLVILLGLCGVLLPGVPGSWLVWAGVLWWALKDPR
                        PLAWAVLVGSTVVLLLSRAVRWALPTRRKRRDEAMRRLTAYAGAGAVLGFVLVPVLGA
                        VPGFMGGIYLAERLRLGRHGEAMASLRTAMRQGGADLLTELFACLLITGAWLGAVLAG
                        "
     gene               complement(14986..15969)
                        /gene="SC1A9.17c"
```

```
     CDS              complement(14986..15969)
                      /gene="SC1A9.17c"
                      /note="SC1A9.17c, possible regulatory protein, len: 424
                      aa; similar to TR:Q56084 (EMBL:D63904) ESA36 protein of
                      undefined function from Streptomyces thermoviolaceus (334
                      aa) fasta scores; opt: 313, z-score: 401.9, E(): 4.1e-15,
                      (30.0% identity in 337 aa overlap). Also similar to
                      SW:BRPA_STRHY bialaphos biosynthetic pathway regulatory
                      protein from Streptomyces hygroscopicus (256 aa), fasta
                      scores; opt: 234, z-score: 233.8, E(): 9.5e-06, (30.1%
                      identity in 163 aa overlap). Contains helix-turn-helix
                      motif from: 1 to: 424, Score 1064 (+2.81 SD). Possible
                      coiled-coil from 171 to 200 (30 residues) Max score: 1.447
                      (probability 0.87)"
                      /codon_start=1
                      /transl_table=11
                      /product="putative regulatory protein"
                      /protein_id="CAA22387.1"
                      /db_xref="GI:4007702"
                      /translation="MLGVLGLEDTHEAAYRALVSVGAADVPDLARRLALGERDTERAL
                      RRLEQNGLAAQSSARPGRWVAAPPGVALGALLTQQRHELERAELAAALLAEEYRAAAA
                      EPAVHDLVEVVTGAGAVAQRFLQLQLGASEEVCALVTDKPVAVTGMENDAEEQATGRG
                      VRYRVVVERSVLDLPTGITELTAALGRDEQVRVVDRVPTKLIVADRSLALVPLTARSS
                      EPAALVVHASGLLELLCGLFEAVWRDALPLRLGASGVTEQAPDGPDGTDLEILSLLLA
                      GLTDASVAKQLDLGLRTVQRRVKRLMELTGVTTRLQLGWHAYERDWVARRD"
     gene             16892..18064
                      /gene="SC1A9.18"
     CDS              16892..18064
                      /gene="SC1A9.18"
                      /note="SC1A9.18, unknown, len: 390aa"
                      /codon_start=1
                      /transl_table=11
                      /product="hypothetical protein SC1A9.18"
                      /protein_id="CAA22388.1"
                      /db_xref="GI:4007703"
                      /translation="MSSPASGAGRPAAPRRPDPRGPTTAGSGAADLREPPEAVPPPRT
                      LPDLPPLPPRFLPSTEADRDYVGSGPPTYDAEPTALPPADPDGLDGLVPDTVLEGARY
                      GTCTLRAVSVRGDSARYRGEPRRDALLVARFGAGEQALVLVAMATGARATAGAHRAAA
                      EVCRWIGRAVGRSHARLAEDLRAARRGDLKSGLHRLTDRSLGRLRAGAAEQGLAPDEY
                      AATLRCLLLPADPGCRTRVFFGVGAGGLLRLRDGAWQDMEPDAGDVTGEPVLGFGSAP
                      RDAPRDAPRETPEDDRLTMDLGITTSPGPYEGPPAGPPREPFRFRASVARPGDVLLMC
                      TAGLAEPLLSEPGLGELLARRWAARPAPGPGEFLADSGVRVKGYADDRTAAAVWEA"
     gene             18093..19835
                      /gene="SC1A9.19"
     CDS              18093..19835
                      /gene="SC1A9.19"
                      /note="SC1A9.19, poxB, pyruvate oxidase, len: 580aa;
                      similar to many including POXB_ECOLI (EMBL:X04105) PoxB,
                      pyruvate oxidase from Escherichia coli (572 aa) fasta
                      scores; opt:2086, z-score: 2458.4, E():0, (53.1% identity
                      in 571 aa overlap). Contains Pfam match to entry PF00205
                      TPP_enzymes, Thiamine pyrophosphate enzymes, score 499.70,
                      E-value 2.4e-154."
                      /codon_start=1
                      /transl_table=11
                      /product="pyruvate dehydrogenase"
                      /protein_id="CAA22389.1"
                      /db_xref="GI:4007704"
                      /translation="MAKQNVAEQFVDILTRAGVERLYGVVGDSLNPVVDAVRRHSGIE
                      WVHVRHEETAAFAAGAEAQITGKLTACAGSCGPGNLHLINGLYDAHRSMAPVLALASQ
                      IPSSEIGLGFFQETHPDQLFRECSHYSELISSPKQMPRLLQTAIQHAVGQGGVSVVSL
                      PGDIADEPAPQGAAETALVTSRPTVRPGDEEIDRLVRMIDDADKVTLFCGSGTAGAHA
                      EVMEFAGKLKAPVGHALRGKEFIQYDNPYDVGMSGLLGYGAAYEATHECDLLLLIGTD
                      FPYNAFLPDDVKIAQIDVRPEHLGRRSKLDLAVWGDARETLRCLIPRVKEKKNRRFLD
                      RMLKKHADALEGVVKAYTRKVDKHVPIHPEYVAALLDEMADDDAVFTVDTGMCNVWAA
```

```
                              RYISPNGRRRIIGSFSHGSMANALPMAIGAQFTDRRRQVVSMSGDGGFTMLMGDFLTL
                              VQHDLPVKIVLFNNSSLGMVELEMLVAGLPSHGVANKNPDFAAVAEACGAFGVRVEKP
                              KDLAGALKAAFKHKGPALVDVVTDPNALSIPPKISADMVTGFALSASKIVLDGGVGRM
                              LQMARSNLRNVPRP"
misc feature                  18132..19679
                              /gene="SC1A9.19"
                              /note="Pfam match to entry PF00205 TPP_enzymes, Thiamine
                              pyrophosphate enzymes, score 499.70, E-value 2.4e-154"
gene                          19979..20386
                              /gene="SC1A9.20"
CDS                           19979..20386
                              /gene="SC1A9.20"
                              /note="SC1A9.20, possible regulatory protein, len: 135aa;
                              similar to TR:O69204 (EMBL:U33059) hypothetical protein
                              from Amycolatopsis mediterranei S699 (144 aa) fasta
                              scores; opt:225, z-score:335.1. E(): 2.2e-11, (41.7%
                              identity in 115 aa overlap). Also similar to TR:Q53897
                              (EMBL:X60316) AbaA regulatory locus for antibiotic
                              production in Streptomyces coelicolor (192 aa), fasta
                              scores; opt: 155, z-score: 284.6, E(): 1.4e-08, (33.3%
                              identity in 120 aa overlap)"
                              /codon_start=1
                              /transl_table=11
                              /product="putative regulatory protein"
                              /protein_id="CAA22390.1"
                              /db_xref="GI:4007705"
                              /translation="MPQLRRRLGRADLRAVPEARRALRELLRHWGGPGQSEVAELLAS
                              ELVTNALVHTDEGAVLTATVGPRALRVEVRDFVGRGRQPRPRAPQREESTNGRGLVLV
                              ESLADDWGVQPCEVGKSVWFELGAGAEAEAEAA"
gene                          complement(20444..21508)
                              /gene="SC1A9.21c"
CDS                           complement(20444..21508)
                              /gene="SC1A9.21c"
                              /note="SC1A9.21c, possible transmembrane protein, len:
                              354aa; contains possible membrane spanning hydrophobic
                              domains in the N-terminal region and possible coiled-coil
                              from 221 to 250 (30 residues) Max score: 1.449
                              (probability 0.87)."
                              /codon_start=1
                              /transl_table=11
                              /product="putative transmembrane protein"
                              /protein_id="CAA22391.1"
                              /db_xref="GI:4007706"
                              /translation="MRLTDISLNWLLPGAVLLLGMLAAVAVLARGKRSSGKDAGADDS
                              WERMEERRRRKEALYGTFSYVLLFCCAAVAAALSFHGLVGFGEQNLGLSDGWQYLVPF
                              GLDGAAMFCSVLAVREASHGDAALGSRILVWAFAFAAAWFNWVHAPRGLGHAGAPHFF
                              AGMSLSAAVLFDRALKQTRRAALREQGLVPRPLPQIRMVRWLRAPRETYRAWSLMLLE
                              GVRSLDEAVEEVRDDRRQKEEKKLRRREQERLERAQLKAISRGHGHRGFPGRGGRQVE
                              VEVQQVERGSERAAAEPAISTPEQLPAASRRPSLQPVRSGSEQMSVGTVDLTAEDDTQ
                              ALPRLDSLERKLKDLEQQFG"
gene                          complement(21766..22611)
                              /gene="SC1A9.22c"
CDS                           complement(21766..22611)
                              /gene="SC1A9.22c"
                              /note="SC1A9.22c, unknown, len: 281aa"
                              /codon_start=1
                              /transl_table=11
                              /product="hypothetical protein SC1A9.22c"
                              /protein_id="CAA22392.1"
                              /db_xref="GI:4007707"
                              /translation="MPVARTALTDAYTRLSEVLPGLGVTELAAADEVPSGGGWVTAAS
                              LAAGGTELAAFLAWDEAQVLRDYGQRARPDVIASFGLHRYAWPACLLITVPWFLHRRV
                              PRHPAAHVSYDRTAAGLPLGRMAVRAASFACLPGDPAAALPGARVVADEEALRAEVRA
                              AVAEHLEPVLAGFGPRMRRRGRALWGMATDEVVEGLWYVAHLLGEQERARHELELLLP
                              GATRPYVGKAAFRELKGPDGAPLHTRDRASCCMFYTLRPEDTCATCPRTCDADRVDKL
```

.

```
                          LAAAG"
     gene                 22780..23601
                          /gene="SC1A9.23"
     CDS                  22780..23601
                          /gene="SC1A9.23"
                          /note="SC1A9.23, possible gntR family transcriptional
                          regulator, len: 273aa; similar to TR:O69996
                          (EMBL:AL022374) WhiH, sporulation transcription factor
                          from Streptomyces coelicolor (295 aa) fasta scores; opt:
                          139, z-score: 267.9, E(): 1.2e-07, (29.8% identity in 292
                          aa overlap). Contains Pfam match to entry PF00392 gntR,
                          Bacterial regulatory proteins, gntR family, score 30.50,
                          E-value 7.1e-08."
                          /codon_start=1
                          /transl_table=11
                          /product="putative gntR family transcriptional regulator"
                          /protein_id="CAA22393.1"
                          /db_xref="GI:4007708"
                          /translation="MKHSAQGAARTGVGTGTGAAAEAVRIPVQPGAADRARGRAPGGA
                          EPDGPARGEHTHGEPPVPRPRALVQRSSVRGQILDALRSALVTGELRPGEVYSAPVLG
                          ERFGVSATPVREAMQQLALEGAVEVVPNRGFRVLERGDRELAELAEVRALIEVPVWLR
                          LARTVPAEHWAELRPLAEGTVRAASSGCPATYAEADRAFHRAALALAGNEQLVRIAGD
                          VHRRAQWPPAGSPSVRGRADLVADAHQHTALLDALIAGDLDVVRALVGDHFTVTS"
     misc feature         23008..23184
                          /gene="SC1A9.23"
                          /note="Pfam match to entry PF00392 gntR, Bacterial
                          regulatory proteins, gntR family, score 30.50, E-value
                          7.1e-08"
     gene                 complement(23847..26234)
                          /gene="SC1A9.24c"
     CDS                  complement(23847..26234)
                          /gene="SC1A9.24c"
                          /note="SC1A9.24c, SecDF, protein-export membrane protein,
                          len: 795aa; similar to many eg. TR:G3220156
                          (EMBL:AF024506) SecDF protein from Bacillus subtilis (737
                          aa) fasta scores; opt:802, z-score:1099.4, E():0, (30.7%
                          identity in 740 aa overlap)."
                          /codon_start=1
                          /transl_table=11
                          /product="putative SecDF protein-export membrane protein"
                          /protein_id="CAA22394.1"
                          /db_xref="GI:4007709"
                          /translation="MKRSRPPSRSRSRTRSRSLNVRALVALAVMAGAVAIALTMPVRL
                          GLDLRGGTQIVLETKSTETTKADREATDRTVEVLRGRIDALGVAEPTIVRSGENRVVV
                          ELPGVQDPKKAADVLGRTAQLTVHSVLGAAEKPADATEGRTETAEDGERVLPDESGQS
                          LRLKAATLTGQDVKGADARFDQQNGAGWTVTVDFKDSGSDRWAQVTGEAACNPAGDPT
                          RRVAIVLDDKIISSPQVDPSVSCGAGITGGSTQITGSFDDAEARELALLIKGGALPVP
                          VETIEQRTIGATLGDEAIDAGAWAAVIGTALTALFIIVVYRLMGALATVALLCYGLIS
                          YAALAAVGATLTLPGLAGFVLAIGMAVDANVLVFERAREEQAARTRPSTRSALTAGFR
                          SAFSAIADSNITTLIAAALLFFLASGPVKGFGVTLGIGVLASMVSALVITRVLAEFAA
                          SRPAVFRRPRITGISSTGPVRDALLRRDPFLMRRPRRWLAASLIVLVVAGSGILVRGL
                          NFGIEFTGGRLIEYSTATQVDPDRARDALADAGFPRAVVQSSGDGDLTVRTEELTDTE
                          AATVTKAVAELGGETEKVRDELIGPSLGEELRRDALIALGLALAAQLAYLAVRFRLLF
                          GTAAVGALAHDVVILVGVFAWLGKPIDGVFLAALLTVIGYSVNDSVVLFDRIRELLGK
                          ERKAPFDRLTNDAILQTLPRTVNTGMGAVLILASLAILADDSLTDFALALLIGVGVGT
                          YSSVFTASPLAIELHNRDTGSRPGRRRGGRTRATGKSEKPGKQVPTSRTERQEVR"
     gene                 complement(26314..26703)
                          /gene="SC1A9.25c"
     CDS                  complement(26314..26703)
                          /gene="SC1A9.25c"
                          /note="SC1A9.25c, probable secreted protein, len: 129aa;
                          Contains a strong signal sequence."
                          /codon_start=1
                          /transl_table=11
                          /product="hypothetical protein SC1A9.25c"
```

```
                             /protein_id="CAA22395.1"
                             /db_xref="GI:4007710"
                             /translation="MRRRLASVLTVLLTVLLPLVPAWPAAGAHVGPSGPLAAAAATAV
                             PHPALDLHADDGCTPVCAAQPRARHDQPAGRPTAPDQHPATTAHLGGGAAPCGHARTS
                             FAPGPVPVSPGRASHDSGRAPPVSSGT"
gene                         complement(26808..27590)
                             /gene="SC1A9.26c"
CDS                          complement(26808..27590)
                             /gene="SC1A9.26c"
                             /note="SC1A9.26c, probable luxR family response regulator,
                             len: 260aa; similar to a family of response regulators eg.
                             TR:O69961 (EMBL:AL022268) probable two-component
                             transcriptional regulator from Streptomyces coelicolor
                             (219 aa) fasta scores; opt:1079, z-score: 1434.4, E():0,
                             (77.5% identity in 218 aa overlap) and DEGU_BACBR
                             (EMBL:L15444) transcriptional activator protein DegU from
                             Bacillus brevis (236 aa) fasta scores; opt:378, z-score:
                             682.5, E(): 9.7e-31, (36.6% identity in 227 aa overlap).
                             Contains PS00622 Bacterial regulatory proteins, luxR
                             family signature, Pfam match to entry PF00196 GerE,
                             Bacterial regulatory proteins, luxR family, score 99.70,
                             E-value 5.6e-26, Pfam match to entry PF00072 response_reg,
                             Response regulator receiver domain, score 132.90, E-value
                             5.6e-36 and an helix-turn-helix motif from: 1 to: 260,
                             Score 1221 (+3.35 SD)."
                             /codon_start=1
                             /transl_table=11
                             /product="putative response regulator"
                             /protein_id="CAA22396.1"
                             /db_xref="GI:4007711"
                             /translation="MSDPSLPEPSEPSARGRHAGASTTPTTPTSASTTSAYGTPPASA
                             PSKIRILLADDHALVRRGVRLILDREPDLEVVAEAGDGAEAIDMARAHEADLAVLDIA
                             MPRLTGLQAARELAALKPGLRILMLTMHDNEQYLFQALKSGACGYVLKSVADRDLVAA
                             CRAAMRDEPFLYPGAVTALIRNYLDRVRHGEETSDHILTPREEEVLKLVAEGHSSKEI
                             AEILFISIKTVQRHRANLLQKLGLRDRLELTRYAIRAGLIEP"
misc feature                complement(26874..26957)
                             /gene="SC1A9.26c"
                             /note="PS00622 Bacterial regulatory proteins, luxR family
                             signature."
misc feature                complement(26952..27293)
                             /gene="SC1A9.26c"
                             /note="Pfam match to entry PF00072 response_reg, Response
                             regulator receiver domain, score 132.90, E-value 5.6e-36"
misc feature                complement(27393..27590)
                             /gene="SC1A9.26c"
                             /note="Pfam match to entry PF00196 GerE, Bacterial
                             regulatory proteins, luxR family, score 99.70, E-value
                             5.6e-26"
gene                         complement(27599..28510)
                             /gene="SC1A9.27c"
CDS                          complement(27599..28510)
                             /gene="SC1A9.27c"
                             /note="SC1A9.27c, probable sensor kinase, len: 303aa;
                             similar to a family of sensor kinases eg. TR:O69960
                             (EMBL:AL022268) probable two-component sensor protein from
                             Streptomyces coelicolor (358 aa) fasta scores; opt:1037,
                             z-score: 1477.8, E():0, (59.2% identity in 299 aa overlap)
                             and TR:Q52558 (EMBL:U02041) a two-component sensor
                             regulating virulence genes of Pseudomonas solanacearum
                             (502 aa) fasta scores; opt: 388, z-score: 475.5,
                             E():3.3e-19, (31.5% identity in 270 aa overlap). Contains
                             hydrophobic N-terminus."
                             /codon_start=1
                             /transl_table=11
                             /product="putative sensor kinase"
```

```
                              /protein_id="CAA22397.1"
                              /db_xref="GI:4007712"
                              /translation="MVLGFATALLLWAPVTVSVPILLTEAVVLVVGMGVMLVANGALL
                              RWGLAPLERLTRLMTTVDLLRPGQRLPISGGGEVPELIRTFNAMLDRLENERATSSAR
                              VLLAQEAERRRIAQELHDEVGQSMTAILLVLGRAADDAEEPLRDELHQAQEITRESLD
                              EVRRLVRRLRPGVLDDLGLISALSSLTHDFATHTGLRVVRRFDADLPVLDHETELVLY
                              RVAQESLTNAARHADAERLEVGLAHADAAVTLTVADDGRGIEAAHEGAGIRGMRERAL
                              LIGAALDITSAPGAGTRIRLTAPLPRK"
     gene                     28890..29216
                              /gene="SC1A9.28c"
     CDS                      28890..29216
                              /gene="SC1A9.28c"
                              /note="SC1A9.28c, unknown, len: 192aa; some similarity to
                              TR:O66611 (EMBL:AE000680) putative protein from Aquifex
                              aeolicus (145 aa) fasta scores; opt:188, z-score: 294.0,
                              E():4.4e-09, (36.6% identity in 93 aa overlap)"
                              /codon_start=1
                              /transl_table=11
                              /product="hypothetical protein SC1A9.28c"
                              /protein_id="CAA22398.1"
                              /db_xref="GI:4007713"
                              /translation="MSLDASRIEPRPERLTAHEARQRLEHARNTRVTQLQALAESGQA
                              DDQLMSAQKAAIERVLKEIDEAFARVEEGTYGACLGCGKPVPGERLEILPYTRYCVAC
                              QRRAAA"
     gene                     29286..29400
                              /gene="SC1A9.29c"
     CDS                      29286..>29400
                              /gene="SC1A9.29c"
                              /note="SC1A9.29c, partial CDS, unknown, len: 38aa"
                              /codon_start=1
                              /transl_table=11
                              /product="hypothetical protein SC1A9.29c"
                              /protein_id="CAA22399.1"
                              /db_xref="GI:4007714"
                              /translation="MNNQIIGDRDTRLPPLSPEDLAALRDNLREQRLFREEQ"
BASE COUNT      4022 a  11280 c  10476 g    3622 t
ORIGIN
        1 gatcacctcg tccacgtcgg cggccagcgc gaaggcctcg tccagcggct gcccggcgcg
       61 ggtgcggtcc ccggccaccg tgtacgacag ccaggcgggc acgtccagtc cgcgcaccgc
      121 cctcagcagc gccgccgcct cgtcggtgtc cgggaccgtc tccagcgcga ggacgtcggg
      181 ccgtgcggcg gccagcacct ccaggcgggg gcggtggaag cgctccagtg cgcccctgcc
      241 gaggccgtac cgtccccggt actcggaacc gtccgcgagc atcgctccgt acgggcccgc
      301 cgacgccgcc acccacagtg cgcgctccgg gcgggccgtc cgcgcccggc gggcggcctc
      361 gcgggccgac gccacgctga gcgcgagcag ttcggcggcc cgctcccgcc cgatcccgcg
      421 ccgggcgaag ccctcgaagg tggcctggta gctggaggtg atcgccacct cggcgcccgc
      481 ctcgaagtag gcgaggtggg cccgcgtgat cgcctcgggg tcctcggcga gcagccgggc
      541 cgaccacagc gcgtcgccca ggtcgtgccc ggccgcctcc agctggttgg acagcccgcc
      601 gtccagcacg agcgggccgg aggcgagggc gtcggcgaag tcgctggtca tggcacgagg
      661 ctacgacctg tccggcgggg ccacgggccc ttgtcccccg agcgagctac agacaggtgt
      721 ccaccgtgag gtgacgatct ccggccgccg gatccgtagc gtccggtgca gcccgcggca
      781 ctcgagccac gggcccgacg aaggagttcc cgtgaggttc gtgtggcagt tcctggccgt
      841 gctggtggcc tacgccatcg gcggcatcgc cgtccaggcg gtgaaggaca acgactggct
      901 cacgctggtg gtcggtctca cgtcggtggc gctggtggtg ttcgtgtacg cctgggtggt
      961 gcggcggacc gagcgccggg aggccctgga cgtggccctg gacgggggccg cggccaaggc
     1021 gggctggggg acgctgatcg gcttcggact gttcggggcc gtcatcacga acctcttcgc
     1081 ctccgggtac tacgaggtcg acggactcgg ctcggtccag ggtgcgatcg ggctggtcgg
     1141 gttcatggcc gccgccgccg cgacggagga ggtcgtgttc cgcgggggtcc tgttccggat
     1201 catcgaggag cacatcggca cctacctcgc gctgggcctg accggcctcg tgttcggcct
     1261 catgcacctg ctcaacgagg acgccaccct gtggggcgcc ctcgccatcg ccatcgaggc
     1321 cggcttcatg ctcgccgccg cgtcgggcgc gcgtcttcagc accgtcgtct ccggcaacgg
     1381 gcacttcggc tggaacttcg ccgcgggcgg cgtcttcagc accgtcgtct ccggcaacgg
     1441 cgacagcgag ggcctgctgg acgccaccat gtcgggtccg aagctgctca ccggcggcga
     1501 cttcggcccg gagggcagcg tgtactcggt gggcttcggg gtcctgctga ccctggtgtt
     1561 cctgtggctg gcgcaccggc gcgggaacat cgtggccttc ggctcgcggc gccgtgccgc
     1621 gggcgccaac tccgccgcta cacttccccg gtgatcgatc gacgacgggt cctggagctg
```

```
1681 tggcgtcgac tcgacgtcac ggtccgggac ctcccgctcg gggtgctgct gctcgtcgcg
1741 tcgctgctgc cgtcgctccg cggtcagggc acggagatcg gcggcctgcc gacccgcccc
1801 gccgacgcgc tggccggggt ggcggccgtc ctccagtcca tccccctggc cgtgcgccgg
1861 cggtggacgc tcctctgcct caccctggtc tcgctcggct tcgccctcga ccaactgcgc
1921 gcctaccacc tgttcgcggg cgccgcgctc ccgatcgtgc tgatcaacgc gggctcccac
1981 caggagaagt accggcgtgc cacgcaggtc accgccaccc tcgggtacgt ggccatggcc
2041 gtcgggctga cgcgcgcgg cggcgacgag acgctggtcg agtacgtgac gttctatctg
2101 gtcctcgccc tggcctgggg catcggcgcg tggatgcgct ccgcgcgggc cgcggaggcc
2161 gaacgccgca gccgggtcgc cgaggacgcc cgcaacgccg aacggacccg catcgcccgc
2221 gagttgcacg acgtcgtgac ccaccacgtg acggcgatgg tcgtgcagtc cgaggcggcc
2281 cggtacctga ccgccgcgcc cgagcgcctc gacgagagtc tggccgcggt cagcgacacc
2341 ggccggcggg ccatcaccga cctgcggcac ctgctcgacc tcctcaaccc cgaccacggc
2401 accgccgagc ccaggacacc acccgtcggc cgggtgctca cgctggtcga gcagacccgc
2461 cgggccgggc agccggtgga gttcaccgag gagggcaccc cggcggcggc caccggcagc
2521 tccgacctcg tggcctaccg tgtcgtccag gaggccctga ccaacgccct caagtacgac
2581 cacggcggca ggacctcggt cctggtgcgg cacggggaac gggagatcac ggtggaggtc
2641 ggcacggacg gctccggctc gggggccgcg tccccccggcg gcagcgggcg ggggctggcg
2701 ggcctgcggg agcgggtcga cgtactgggc ggcgagttca gcaccgaccg tccggcggac
2761 ggcggcttcg tggtccgggc ccggataccc ggcgggagcg ggggaagcac ggcatgagcg
2821 cgccgatccg ggtcgtgatc tgcgacgacc aggcactcat ccgcaccggg ctcgcgacga
2881 tcgtcgacgc ccagcccgac ctcgaggtgg tcggcgagtg cggggacggg cagaccgggg
2941 tcgacctggc ccgcgaactg cggccggacg tcgtggtgat ggacatccgc atgccggtgc
3001 tcgacggcct ggaggccacc cgcctgctcg ccggtgccgg agtggcgcac cccgtcaagg
3061 tgctcgtggt gacgacgttc aacctggacg agtacgtcta cgaggcgttg cgcgcgggcg
3121 ccagcggctt cctgctcaag gacgcgcccc cggaccggct gctgcacggc atccgcaccg
3181 tggccatggg cgcggcgctg ctcgaccccg atgtgacgcg ccgcctcgtg ggccggtacg
3241 cggcccggat ccgccccgcc gagggcaccg cgcgggacat tccgctgacc ccccgggaga
3301 cggaggtcct gcgcctgatc gcggacgggc tgtccaacag cgagatcgcc gcggccctcg
3361 tgatcagccc cgagaccgtc aagaccttcg tctcccgcat cctcaccaaa ctcgacctgc
3421 gcgaccgcgt ccaggcggtc gtcttcgcct accgccacgg gctggtgacc tgacgcccgg
3481 ccgccggccg ctcagccccg caaccggccg cgcatgtgcg acagacgccg gtcgatgcgc
3541 aggttgtacc aggccagcca gcccaggaac acgaccccga ggccgagcat caggcttccg
3601 gcaccgaccg agccggagag gtaccacagg gccgaggtgc cgaagaagat gacggccagc
3661 atcgggaacg cccaccagcg gttgcgccgc agccgctcct ggcgcgagtc gacgaacgcc
3721 gccatcgccc gccgccgctc cgggtcccgg ggcgcgggcc cgcccttgag gatctgccgc
3781 tccatggcgg gcacgtcgtc cggttccgca ccgaccgccc tggcgtcgcg cacgcgccgg
3841 cggcgcaccc accacacgat ccagacggcg gccggcaccg cgcccagcag agccgacagc
3901 cacgccgacg tgccgttcag gaccaaccgg accacgacgc cggccagcag ggccatcact
3961 ccgatgcccc acaggctgtc ccgcctgcgc agccgcggac ggtcgtcgct tctctctgtg
4021 ctcgccatgt cacctctcat ggggaagacg tggacgtacg gtgcagaccg gatacccgcg
4081 ttcgcccgga acacgtcgcc gggacacttc ggccggcgcc gaagcaaccg gcccatagcg
4141 tgcgggacat gaacgacacg cgaaccccca cgagcaccgt caccagcacc cgcacgcccg
4201 ccgacgccgt caccggcatg gtcgaccacg tcctgcacct ggccgccggc tggacccgct
4261 gggacgggac gcccgcccac gtcgacggcc gcgtctacac cccgcacaag gcgatccggc
4321 gggtcgccga ccacctggtc gaccacctcg cggagctgga ggcccggctg gcggggggagg
4381 agacgcagcc cgaccactgg cacgcctccc tggtcaccac ggaagcggac cgcgccgcct
4441 tcaccgccga ggatctcgac gaggcccgca gccgcctcac ccggctcgcc cggatctggg
4501 ccaaccgcct cgacgcgctc accgacgagc agctcgacca ctcgccgggc gagggctgga
4561 gctttcgcga actggccgcg cacctcgcgg agtccgggta ctacgccgac gcggtgggcg
4621 acctgtcgtg acagcgaccg ccttcgccgc gctgcaccgc gccggtgagc ccctgctgct
4681 gccgtgcgcc tgggaccacg cctccgcctt cgccctcgcc gggcagggct tccgcgccgt
4741 cggcacgacc agcctcggtg tcgcggcggc ggccggactg cccgacgggg cgtccgcgac
4801 ccgggacgag acgctgcggc tggccctcgt cctcggctcg gcccccttcc cgctctccgt
4861 cgacgcggag gacgggttca gcgacgaccc cgacgaggtg ggggagttcg cgcggcagct
4921 ggcggccgtc ggcgccgtcg ggatcaacct ggaggacggt ctcggcccg tcggccggca
4981 cgccgcgaag atcgccgcgg tccgctcggc cgccccccggc ctgttcgtca cgcccgcac
5041 ggacacctac tggtcgggcg acggcgacgt ctgcggcgcc tggaggccta
5101 ccgcgaggcc ggtgccgacg gcgtcttcgt ccccgggctc accgaccccg cccgcatcgg
5161 gtccctcgcc gcccgcttcg acgtgccact caacgtcctg tacacgcccg ccggacccgg
5221 cctcgcgcac ctcgccgacc tcggcgtacg ccgcgtcagc ctcggctcgc tgctctaccg
5281 gcggcgctg ggcgccgcgc tgcgggcggc cgccgacgtc cgggcggggcc gggacccggg
5341 cggtccgacg cccacgtacg acgaggtgcg ggcgccgggc tgacctcagg cgcgcggcca
5401 gttctccagc gccacgtgca gcgcgtcgac cgccttgtcc caggacgccc gcacgtcccg
5461 cggggcgccg aaggcgccgg aggactccag cgcgcagtac ccgtggaagg tgctgcgcag
5521 caggcgcacg gcgtcggtga ggtcgggctc ggtgaggccg taggaccgca gcatgccgta
```

```
5581 ggtgatctcg gcggtgcggc gcagcgcggc ggagtcggcg accagggact ggtcgacgcg
5641 gatctgcgtg gcggcgtagc ggccggggtg gcgcagggcg tacgcccggt aggcgcccgc
5701 gaaggcggcc agcgcctcct tcccggcccg gccggccacc gctccggcga tctcctcgat
5761 cagctcgccg ccggccagca gcgcgaccct gacgcgcagg tcctggaggt tgcgcacgtg
5821 cgtgtacagg ctggcgtcct tcaccccgaa gtggcgggcg agcgcggaga cggtgaccga
5881 ctcgaagccc gtctcgtcgg ccaggtcggc ggccgcggcg acgacgcggt ccgtggtgag
5941 acccactcgg ggcatgggcc gccttcccgc aacctggagt tcctaggttt catgctaggg
6001 caggacgggc ggcagccgcg gatcacccgc gagtcgttca acgttcacct tccccgccga
6061 ttctcctcga agccgcgcac ggcggcggtc ttcaggagga acacgcatgg ggcacggaca
6121 cggacacccg cacggtcacc accaccacgg acatacgcac gaccacgagc acgccaccga
6181 acaggtcctg cccgccgcct tcgacaccag cgtccccgac gaggccctga gccccgcgca
6241 gcagtcgcgc cggggcctgc tgcgccgtgc cggcctgctc ggcgccggac tggccgccgg
6301 taccgtcctc gccccggccg ccaccgccac cgccgccccc gcgcgggccg cgagcaacgg
6361 ccgccgcggg aagggcttcc tctggctggc cggcgaccac cacatccaca cccagtacag
6421 cagcgacggc aagtaccgcg tcgtcgacca ggtccgccag ggcgcccggc acggcatgga
6481 ctggctggtc atcaccgacc acggcagcaa cacccacgcc aagatcggtg tggagaaggt
6541 caaccccgac atccgcgagg cccgcgccgc ccacgaggac accctcgtct tccagggcct
6601 ggagtggaac atcccggccg ccgagcacgg caccgtcttc gtccaccccg gcaagcacga
6661 ggtctccgtc ctcaagcagt tcgagaccga ctacgacggc agcgtcaagg cgcgggcga
6721 ctccaccccc gccaacgagg ccctcgccat cgccggactc tccttcctcg ccgaccaggt
6781 caagcggcgc aaggtcaagg acgccctgat gctcgccaac cacccggcgc gcaagggcat
6841 cgactccccg cacgagatcc gcgcctggcg cgacgccacc tcccgcggcc accagatcgc
6901 ggtcggcttc gagggcgccc cgggccacca ggcgggcggc ctgcccgagc ccctcggccc
6961 gggcggcgcc cgcggcatct acgacggcag cccgagcgcc aactccttcg ccggctaccc
7021 cctggagagc taccgcacct ggggcggctt cgactggatg accgccaccg tcggcggcct
7081 gtgggacagc ctcatcgccg agggccgccc ctggtggatc accgccaact ccgactccca
7141 ccaggtctac gccgacaccg gcgcccgcgg cggcggcgac ttcaacgcca acggccgcta
7201 cgacgacccg gtctacgcgg ggcagatcga catcacccag aacgacttct ggcccggcca
7261 gtacagccgc acccacgtcg gctccgacgg cttctcctac gccgccgtca tggacggcat
7321 ccgggccggc cgcgtctggg tcgaccacgg gcagctcgtc agcggcctcg acgtccgcgt
7381 ctccggcggc ggccgctggg ccacgctcgg cggcgccctg cacgtccgca ggggcacccg
7441 ggtcaccctc tccatcgacg tcgccctggc cggcggcccc aactgggccg gattcgtgcc
7501 gaagctggcc cgcgtcgacg tcatccaggg cgacgtgacc ggaccggccg ccgacaagga
7561 caccttcacc gcgccgaccg cccgggtcgt caagtcctac gaggtcgaca aggaaaccgg
7621 caccgtccgc ctcacctacg acctcggccg cgtcgaccgc cccgtctacc tccgcacccg
7681 cggcaccgac ggcaaccggt ccgccgtcgg ctcgctgggc gccaaggtcg acccggccgg
7741 ccccgccatc gacgtcgtcg gcgacgccga cccgtggcgc gacctgtggt tctactccaa
7801 cccggtctgg gtcctgccct cgtgaccccg cccgccgcgc cgtacgtcct cggcatcgac
7861 accgacgcga cgaccctgcg cgaggccgat cacctgctcc aggcgctggc ggccgaactc
7921 gacctccccg agggcgtctt cggctgcacc cacctggtgc gggacggccg cccgcgggtc
7981 gccctgtccc tcgccgccga ggcggagccg gtcctgcgca ccgcccggga ccggctcacg
8041 gcccggggcc acgaggtccg cgacggcacc tgggacgaga gcggccgcgc ggtcctcttc
8101 cccggtgcgg ccgccctcac cggcaccctg accctcgccg agctgctcgc ccgctccgcc
8161 gtcgaccggg tgaccgtcct cggcacaccg gacgagccga gcccggacac ccggctggtc
8221 acccggaacc acgtccgccc gcactggcag gacggccgcc tcgtgctggc cgccatgccc
8281 gcggtcggcg gcaccctcgt accgttcgag gaccccgacc cgacaccctg ctgcgccgac
8341 cactgacccg gcccggcgcc gaggggcgcg gggtgctcag cccgcgagca ccccgcgcac
8401 gaacgcgttg gtgaacttcc ccgccgggtc cagcgccccg gccagcgccc cgaagtcggc
8461 cagccgcggg tacagcgcac gcagctcgcc cgccgggacg gtgaacacct tccccagtg
8521 cggacgggcc gcgaagggga cgagcgcctc ctccagccgc cgcaccaccg gcagcaccgc
8581 cgccgtgtcc tcgacccagg tgaagtgcgc ggccacggtg tcccgcccgt acgccgggct
8641 cagccactgc gcgtcggcgg cgaccgtgcg gatctcgcag gtctggagca ccggcgcgag
8701 cgtctcccgt atcgcgtcca tcgcgtgcag ggcggccagg gcgtgctccc ggggcatcag
8761 gtactccgac tgcaactcgg caccgctgct gggcgtgaac tccgcgcgga agtgcggcag
8821 ccgctcgtgc cagggccccg gcaccccgaa ctgctccgtg cagttcaccg cgggcatgcc
8881 cggcaccgga tgcatcttct cggcggccgg ggccgcgtac gggaaaccgt ccagcggccg
8941 gtcggtgcgc cgcttcagcc acacctgccg gaaaccgggc gcccgccagt cggtgaacag
9001 actcacgctg tacgccgccg ccatcaccgt ctcgaacgtc gccgggtcca acccggccag
9061 cggcagctcg gtgaagacgt gctgttccat ctcgtacgcc ggctccaggt cgagtgtgag
9121 cgacgtcacc acgcccagcg cgccgagcga ggtcaccgcc ccgccgaacc gctcgtcgcc
9181 ccgcgcgatc accacggtcg aaccgtccgc ggtgaccagc tccacctcgc gcaccaccga
9241 ggccagcgaa ccgttgccca cccccgaacc gtgggtgccg gtggccaccg acccggcgac
9301 cgagatgtgc ggcagcgagg ccatgttcgg cagcgccagg ccccgcgcgt gcaccacccg
9361 ggccagctcc gcgtaccgca caccgccgcc gacccgcacc gtacgggccg ccgtgtccac
9421 gtccaccacg gacggcaggc ccgccagcga cagcaggaca cccccgtcgc ccggctcggc
```

```
 9481 gatctcgttg aaggagtgcc cgctgcccag cacccgcacc ctggcgctgt ccgccaccag
 9541 ggcccgcagc gcgtccaggg agtgcggccg cagcagttcc ttcgccgtgt acgtgatgtt
 9601 gccggcccag ttggtgaccg tgatgtcgct catagcgcgg aaacctaccc ggcgcgtctg
 9661 gagggccccc gcgcgggaag cctctcccgc gggggcatac cgtgagaagt cgaacgccgg
 9721 acccgggagg agacacggga tggacagccg taccgcgctg gtcgaggatc tgatggagcg
 9781 gttcccgcac gtaccgcgcg aggccgtctt caaggaggac ctgctgcgcg gcggtgtcgc
 9841 cttcgacccc tccgcactca gtgataccac caacgaggca gccggcgagg tcaagccgaa
 9901 gtcgtacttc atcttctcct tcgaccacgg caccctgccc gagttgggcg aggccgccct
 9961 gcgccgcccg cccgaggaga tcatcctcac cggcggcccc tacgacctgc cccgcaccgt
10021 cgtctcggtg cgcgtgaacc cggcctcgcc ctaccgggtc gccgcgaacg aggacggcgt
10081 cctcggcctc tacctcgacg gcaagcggat cgccgacgtc ggcgtgccgc ccatgcccga
10141 gtactaccgg cacaagctct ccaacgggaa gtcggtcatg gaggtggccc cgaccatcca
10201 gtggggctac ctgatctacc tgaccgcgtt ccgcgtctgc cagtacttcg cgccaagga
10261 ggagtgccag tactgcgaca tcaaccacaa ctggcgccag cacaaggcgg ccgggcggcc
10321 gtacaccggt gtgaaggacg tcgacgaggt cctcgaagcg ctggagatca tcgacaagta
10381 cgacaccgcc aagatctcca ccgcgtacac gctcaccggc ggcgcgatca cgtcgaaggt
10441 ccagggtctg gacgaggccg acttctacgg ccggtacgcc aaggccatcg aggagcactt
10501 ccccggccgc tggatcggca aggtcgtcgc ccaggccctg cccaagccgg acgtccagcg
10561 cttcaaggac tacggcgtgc agatctacca ccccaacttc gaggtgtggg acgagtacct
10621 gttcaagatg tactgccccg gcaaggagcg ctacgtcggc cgcgacgagt ggcacaagcg
10681 catcctggac tccaccgagg tcttcggcgc ccgcaacgtc atccccaact cgtcgcgggg
10741 cgtcgagatg gccgagccct tcggcttcaa gaccgtcgac gaggccatcg agtccaccac
10801 cgagggcctg cgcttcttca tgtcgcacgg catcacgccc cgcttcacca cctggtgccc
10861 cgagcccacc accccgctcg gcaagaccaa ccccgacggc gctccgctgg agtaccacat
10921 ccgcctgctc caggcctacc gccagaccat ggaggactac ggtctctcct cgccccccgg
10981 ctacggcccg cccggcgcag gcaacgcggt cttctcggtc agctccttca tggacagcct
11041 gccggaggac gcgcccgtgg aggtgtgagc ggacggttcc ggggtttatt cggtggcgcg
11101 gccttcgggc gcgcgactag ggtgaccgat gtccgagcgg cggcccaagc ggccgccgac
11161 agggtctttt ccgaggagg tgtcgaccaa tggctgtcac tgtgacgggc gctgcccgca
11221 tccaggagtc catcgagtcc acctccgcgg ccaccggctg atctcagccg ctccctgagt
11281 tctcgcgccg gggccgcacc ccgaagggtc accccttgac ttccacgtct tcccactcca
11341 accactccgc cctgtcctcc tacggctggg acgactcctg ggccgacgcg ttcgcgccgt
11401 acgccgccga agggctgctg cccggccggg tcgtgcgggt cgaccgcgga cagtgcgacg
11461 tcgtcaccgc ggacggcgtc ctgcgcgccg acaccgcgtt cgtcacgccg cacgacccgc
11521 tgcgcgtcgt ctgcaccggc gactgggtcg ccgtcgaacc gggcggcaac ccgcgctacg
11581 tacgcacgta tctgccccgc cgtaccgcct tcgtgcgctc cacctcgtcc aagcggtccg
11641 aggggcagat cctcgcggcc aacgtcgacc acgccgtcgt cgccgtgtcc ctcgcggtcg
11701 agctggacct cgcccggatc gaacggttcc tcgccctcgc ctgggaatcc ggggcgcagc
11761 ccctggtcgt cctcaccaag gccgacctcg tgcccgaccc ggtgaccctg gcgtacctcg
11821 tccaggacgt ggagaccgcc gcgcccggcg tgccggtgct gccggtcagc gccgagcagg
11881 gcgaggggct cgacgtgctc gccgccgtcg tctccggcgg taccgccgtg ctgctcgggc
11941 agtccggcgc cggcaagtcc accctcgcca acgcgctgct cggcgaggcc gccatggacg
12001 tgcaggccat ccgcgacgtc gacggcaagg gccggcacac caccaccacc cgcaacctcc
12061 tcgccctgcc cggcggggggc gtgctgatcg acaccccccgg actgcggggc gtcgggctct
12121 tcgacgcggg caacgggtg accaggtgt cgccgagat cgcggagctg gccgaggagt
12181 gccggttcca cgactgcgcc cacgagagcg agccgggctg tgccgttctc gccgccatcg
12241 acagcggtgc actgcccgag cggcggctgg agagctaccg caagctgatg cgggagaacc
12301 agcgcatcgt cgccaagacc gacgcccggg cgcgggccga gatccgcaag gagtacaagc
12361 gccggggcgc catcggcaag gcggcgatgg aggccaagcg gggtggcctg aggtagcccg
12421 ccggccgctc caacgcgatg tccgatttcc gccagcccgg tccgccggtc tggcggagac
12481 tggacagcgt gacaccacag acagtgcagc cggcagagca cgcggacgcg cgcgaggacg
12541 tccggtacga ggcagtgcgg agccgggacg cccggttcga cggcgcgttc ttcttcgccg
12601 tcgagacgac cggcatctac tgccggccca gctgccccgc cgtcacaccg aagcggcgca
12661 acgtgcggtt cttcgcgacg gccgccgccg cgcagggctc gggcttccgg gcctgccggc
12721 ggtgccgccc cgacgccgtg ccgggctccg ccgactggaa cgtccgcgcc gacgtcgtcg
12781 gccgggccat gcggctgatc ggcgacggcg tggtggaccg cgaggcgtc gccggactcg
12841 ccggacggct cggctacagc gcccgccagg tgcagcgcca gctcaccgcc gaggtgggcg
12901 ccggacccgt ggccctcgcc cgcgcccagc gggcgcacac cgcgcgggtc ctgctgcaga
12961 ccaccgtcct gccggtcacc gagatcgcct tcgcgtccgg gttcgccagc gtgcggcagt
13021 tcaacgacac gatccggggcc gtgtacgcgg ccacccccgag cgagctgcgc gccgccgccc
13081 cggcccgcga ccgggccgcc cggcgcacgg ccaccccctc gccggcgtc ccgctgcgcc
13141 tgcccaccg cggcccctac caggccggcc ccgtcttcga cctgctccag cgggaggccg
13201 tcaccggcgt cgaggaggtg agcggcgaga ccggccgccg cctctaccgg cgcacctgc
13261 gcctgccgta cggcaccggg atcgtcgccg tccaggaacg gccgggccgg gcgggcaccg
13321 gcagcggggg ctggctggag gcccgcctcc acctcaccga cctgcgcgac ctgaccacct
```

```
13381 ccgtacagcg gctgcggcgg ctgttcgacc tcgacgccga cccgtacgcc gtcgacgagc
13441 ggctcggcgc cgacccccgg ctcgcccccc tggtcgccgc ccgccccggg ctgcgctcac
13501 cgggcaccgc cgaccccgcg gaactggcgg tgcgcgccct ggtggggcgc accgaggccg
13561 agcggctcgt ccagcggtac ggcaaggcgc tcgacgcgcc ctgcggcacc ctcacccacc
13621 tcttccccga accggacgtc ctcgccgggg ccgcgccgca cggcaccccg ggcgcgctcg
13681 ccgccgccct cgccgacggc gccgtacgcc tggaccccgg cgccgaccgg gacgacgccg
13741 agcgcgcgct gctcgccgtg cccggcctgg acgcccgtac cgtcgccgtc gtgcgcaccc
13801 gcgccctcgg cgacccggac gtcgccccgc ccggcgccgc cgtccccgac acctggcgcc
13861 cctggcgctc gtacgcactg aaccacctgc gcgcagcagg agagtgggag aacgaccgat
13921 gaccacgacc acgccgacga ccacgaccac gagcatcccc gccgagacgt actggcacga
13981 ggtcgacagc cccgtcgggc cgctgctcct caccgccggc tccgacggcg cgctcacctc
14041 cctgtccgtg cccgggcaga agggcggccg gagcgtccgg gacgggtggc ggcacgacgc
14101 cgggcccttc cgggtggccg aggaacagct cggcgcctac ttcgccgggg agctgacgga
14161 gttctcgctg ccgctgcgcg cgcagggcac cgccttccgg gagcgcgtgt gggcggcatt
14221 ggacgacgtg ccctacggcg cgaccacgac gtacggcgag atcgccgcgc gcatcggggc
14281 gtcgcggccc gcggtacgcg ccgtcggggg tgcgatcggc gcgaatccgc tgctgatcct
14341 gcgcccctgc caccgggtga tcggcgccga cggctccctg accgggtacg cgggcgggct
14401 ggaacgcaag acgcggctgc tgtcgctgga gggcgccccg ctcagccggc cagtaccgct
14461 cccagccacg ccccggtgat cagcaggcag gcgaacagct cggtcagcag gtcggcgccg
14521 ccctggcgca tcgccgtccg cagcgacgcc atcgcctcgc cgtgccgccc cagccgcagc
14581 cgctcggcga ggtagatgcc gcccatgaag ccggggaccg cgccgagcac ggggaccagg
14641 acgaagccga ggacggcgcc cgcgcccgcg tacgccgtca gccggcgcat ggcctcgtcc
14701 cggcgtttcc tgcgggtcgg cagcgcccag cgcaccgccc gggacagcag cagcacgacg
14761 gtggaaccga ccaggaccgc ccacgcgagc ggccgcgggt ccttcagcgc ccaccacagc
14821 accccggccc acaccagcca cgatcccggc acgccgggca gcaggacgcc gcacagcccg
14881 agcagaatga ccaggccggc cagcaggagg tcccacgctc ccatctgtcc agggtgcagg
14941 agtccggcca ggtccgcaga agtccggccc ggcgggcggg gcccctcagt ccctcctggc
15001 cacccagtcc cgctcgtagg cgtgccagcc cagctggagc cgggtggtga ccccggtcag
15061 ctccatcagg cgtttcaccc gtcgctgcac ggtcctgagg cccaggtcga gctgcttggc
15121 cacgctcgcg tcggtcagcc cggcgagcag caggagagg atctccagat cggtgccgtc
15181 ggggccgtcc ggggcctgct cggtcacgcc ggaggcaccc agacgcagcg ggagggcgtc
15241 ccgccacacg gcctcgaaca ggccacacag cagttcgagc agaccgctgg cgtgcaccac
15301 cagcgcggcg ggttccgagg agcgcgcggt gagcgggacg agggcgaggg agcggtcggc
15361 cacgatcagc ttggtcggca cccggtccac gacccggacc tgctcgtcgc ggcccagcgc
15421 ggcggtcagc tcggtgatgc cggtcggcag gtcgaggacc gaccgctcga cgaccacccg
15481 gtagcggaca ccgcgcccgg tggcctgctc ctccgcgtcg ttctccatgc cggtcacggc
15541 caccggcttg tcggtgacca gcgcgcacac ctcctcgctc gcgccgagct ggagctggag
15601 gaagcgctgg gcgacggccc cggcgccggt caccacctcc accaggtcgt gcaccgcggg
15661 ctcggccgcc gccgcccggt actcctccgc cagcagcgcc gccgccagct ccgcccgctc
15721 cagctcgtgc cgctgctggg tgagcagggc gcccagcgcg acgcccggcg ggccgcgac
15781 ccagcgtccc gggcgggccg aggactgggc ggcgagcccg ttctgctcca gccggcgcag
15841 cgcgcgctcg gtgtcccgtt cgccgagcgc cagccgcctg gccagatccg gtacgtcggc
15901 cgcgcccacc gacaccagcg cccggtacgc cgcctcgtgc gtgtcctcca gccccagcac
15961 tcccagcatc aggtgtcgcc cctccccaga accatgccca cctccacctg cgcggacgtc
16021 cgtggcggga aacggccacg gcgcaaaccc gcctcgggac atcatcgccg taccgccggt
16081 cactctgcca aggtggcgtc gccgggtggt cctcccgtgg ggggtggggc cgcccggcca
16141 ccccacctcg gagaactgtt cgacgcgtcc cccgacacgc cgccgggcgc acgccggtcc
16201 gccccgaatt tgccgatgcc ccgtttccat ccggccccat cggtggacaa tcaggagcat
16261 gagtcagcag gggggaaggc ccaccggtcc cgccggtccc accggacctg aggacgactg
16321 gtgggggcag ctgtacgacg actccaccga cgacaccggc cccacggccg cacccgattc
16381 cctggacgac cgcttcgcct cggcgtcgga cacactggcc cgcggaccgg cggccccgg
16441 ccccggcgag cccacgcccc cggggccgcc accgtccccc accgcgcccg aaccgggacg
16501 gccctggtgg cccaggaccc cgcccgccgc cgtccccggc ccccgcaccg gggacgacgc
16561 cccaccgctc cacgcggacc tgccgggacc ggacgacggc gacccgaggc acgacgcccc
16621 accgccgccc gccggggga ccggctcggg gagacgcggg ccgtaccggt ctccgggctc
16681 cccggaggcg ggggacggcg ggccggcga gggtgggccc ccgcgggctg ccgggccgtc
16741 ccggcccacg gctccctggg agcctccgcc cgcgaccccc gccggcccga ggagctttcc
16801 ggccgtgagc gccgaccccg gcgaggcgc gtcggacgcg ccgcggcccg cgcccgccgg
16861 aactccgccg ccgggtgaag cgtcgggtcc cgtgtcgtcg ccggcgtccg gtgccgggcg
16921 gcctgccgcg ccgaggcggc cggatccgcg tgggccgacc acggcgggt ccggggccgc
16981 cgacctcaga gagccgcccg aggcggtccc tccgccccgg accctgcccg acctgccgcc
17041 gcttccgccg cgcttcctcc cctcgacgga ggccgaccgc gactacgtcg gcagcgggcc
17101 gcccacctac gacgccgagc ccaccgcgct gccgcccgcc gatcccgacg gcctcgacgg
17161 cctggtgccg gacaccgtcc tggagggcgc ccggtacggg acgtgcacgc tgcgggccgt
17221 ctccgtgcgc ggggactccg cgcgctaccg gggcgagccg cgtcgcgacg cgctgctcgt
```

```
17281 cgcccggttc ggcgccggag agcaggcgct ggtcctggtg gcgatggcga ccggcgcccg
17341 tgccaccgcc ggggcgcacc gggcggcggc cgaggtgtgc cggtggatcg gacgggccgt
17401 gggccgcagc cacgcccggc tcgccgagga cctgcgggcc gcccggcgcg gcgacctgaa
17461 gtccggcctg caccgcctca ccgaccgcag cctcggccgg ctccgggccg gtgctgccga
17521 acagggcctg gcccccgacg agtacgcggc caccctgcgc tgcctgctgc tgcccgccga
17581 ccccgggtgc cgtacgcggg tgttcttcgg cgtcggcgcg ggcggactgc tcagactgcg
17641 cgacggcgcc tggcaggaca tggagccgga cgccggtgac gtcacgggag agcccgtcct
17701 cggcttcggc tccgcacccc gcgacgcacc ccgcgacgca ccgcgcgaga cgcccgagga
17761 cgaccggctg accatggacc tcggcatcac cacatccccg ggcccgtacg aggggccgcc
17821 cgccgggccg ccccgcgagc ccttccggtt ccgggcctcc gtagcccgcc cgggtgatgt
17881 gctcctgatg tgcaccgcgg gcctcgccga gccactgctg agcgagcccg gactgggcga
17941 actcctcgcc cggcgatggg ccgcccgccc ggcgcccgga ccggggagt tcctcgccga
18001 cagcggcgtg cgggtcaagg gatacgcgga cgaccgtacc gccgccgccg tttgggaggc
18061 gtgagaccgc cgcgctgtggc atagctggac ccatggccaa acagaacgtc gcggaacagt
18121 tcgtcgacat cctcacccgc gccggggtcg agcgcctcta cggagtcgtc ggggacagcc
18181 tcaacccccgt cgtggacgcc gtgcgccgcc attccggcat cgaatgggtg cacgtacggc
18241 acgaggagac cgccgccttc gccgccgggg cggaggcgca gatcaccggg aagctcaccg
18301 cgtgcgccgg ctcctgcggc cccggcaacc tccacctcat caacgggctc tacgacgccc
18361 accgctccat ggcccccgtc ctcgccctgg cctcccagat cccgtccagc gagatcgggc
18421 tcggcttctt ccaggagacc caccccgacc agctgttccg cgagtgcagt cactacagcg
18481 agctgatctc cagcccgaag cagatgcccc ggctgctcca gaccgccatc caacacgccg
18541 tcggccaggg cggcgtcagc gtcgtctcgc tgcccggcga catcgcggac gagcccgccc
18601 cgcaggggc cgccgagacg gccctcgtca cctcccggcc caccgtccgg cccggcgacg
18661 aggagatcga ccggctggtc cggatgatcg acgacgccga caaggtcacc ctgttctgcg
18721 gcagcggcac cgcgggcgcg cacgccgagg tcatggagtt cgccgggaag ctcaaggcac
18781 cggtcgggca cgccctgcgg ggcaaggagt tcatccagta cgacaacccg tacgacgtcg
18841 gcatgagcgg actgctcggc tacggcgccg cctacgaggc cacccacgag tgcgacctgc
18901 tgctcctgat cggcaccgac ttcccgtaca acgccttcct cccggacgac gtgaagatcg
18961 cccagatcga cgtgcggccc gagcacctgg gacggcgctc caagctggac ctcgcggtgt
19021 ggggcgacgc gcgggagacg ctgcgctgcc tgatcccccg ggtcaaggag aagaagaacc
19081 gccgcttcct ggaccggatg ctgaagaagc acgccgacgc cctggaaggc gtcgtcaagg
19141 cgtacacccg caaggtcgac aagcacgtcc cgatccaccc cgagtacgtg gccgccctgc
19201 tggacgagat ggccgacgac gacgcggtgt tcaccgtcga caccgggatg tgcaacgtct
19261 gggcggctcg ttacatctcg cccaacgggc gccgccgcat catcggttcc ttctcccacg
19321 gttcgatggc gaacgcgctg ccgatggcga tcggcgccca gttcaccgac cggcgccgac
19381 aggtcgtgtc gatgtccggc gacggcgggt tcaccatgct gatgggcgac ttcctcaccc
19441 tcgtccagca cgacctgccg gtgaagatcg tgctcttcaa caactcctcg ctcggcatgg
19501 tcgagttgga gatgctcgtc gcgggcctgc cctcgcacgg ggtggccaac aagaacccgg
19561 acttcgccgc cgtcgccgag gcctgcggcg ccttcggcgt ccgggtggag aagcccaagg
19621 acctggccgg ggcgctgaag gcggcgttca agcacaaggg tcccgccctc gtcgacgtgg
19681 tcaccgaccc caacgccctg tccatcccgc cgaagatcag cgccgacatg gtcaccggct
19741 tcgccctgtc cgcctcgaag atcgtcctgg acggcggggt cggacgcatg ctccagatgg
19801 cccggtcgaa cctgcgcaac gtcccccgcc cgtgaccgta cggacagccg gttcgaaacc
19861 ggatgactgg cgagggctcg aacgggcaag gctttccccg tgaagttgtt cgaccggagc
19921 gggacggacg gggggagacg acatcagcgc gccgacgggg gacatgaagg agcaggcggt
19981 gccacagctc aggcggcggc tcggccgggc ggacctgcgg gcggtgcccg aggcccggcg
20041 ggcgctgcgt gagctgctgc ggcactgggg cgggcccggc cagtcggagg tcgcggagct
20101 gctcgccagc gagctggtca ccaacgcgct cgtgcacacc gacgagggcg cggtcctgac
20161 cgcgacggtc gggccgcgcg ccctgcgggt cgaggtacgg gacttcgtgg gccgggggcg
20221 gcagccccgg ccgcgcgcac cgcagcggga ggagagcacc aacggcagag ggctggtcct
20281 ggtcgagtcg ctcgccgacg actggggcgt acagccctgc gaggtgggca agtcggtctg
20341 gttcgaactg ggcgcgggag cggaagcgga agcggaagcg gcctgaggcg ccgggaagac
20401 cggcgcctca ggccatcggt atcggatacg gggcgcgtcc ggctcagccg aactgctgct
20461 ccaggtcctt gagcttgcgc tccaggagt ccaggcgggg cagggcctgg gtgtcgtcct
20521 ccgcggtgag gtcgacggtc cccacggaca tctgctcaga cccgctccgc acgggctgca
20581 gggagggccg ccggctcgcc gcgggcagtt gctccggggt cgatatggca ggctccgcgg
20641 cggcgcgttc cgagcctcgc tcgacctgct gcacctcgac ctcgacctga cggccgccgc
20701 gcccggggaa gccgcggtgg ccgtggcccc ggctgatcgc cttgagctgc gcccgctcca
20761 gccgctcctg ctcgcgccgc cggagcttct tctcctcctt ctggcgccgg tcgtcgcgca
20821 cctcctcgac cgcctcgtcc aggctgcgca cgccctccag gagcatcagc gaccaggcgc
20881 ggtacgtctc gcggggcgcg cgcagccagc ggaccatgcg gatctgcggc agcggacgcg
20941 gcaccagacc ctgctcgcgc agggcggccc ggcgggtctg cttcagggcc cggtcgaaca
21001 gcaccgccgc cgacagggac atgcccgcga agaagtgcgg ggcacccgcg tgccccaggc
21061 cgcgcggcgc gtgcacccag ttgaaccagg ccgcggcgaa ggcgaaggcc cacacgagta
21121 tccgcgagcc gagcgccgcg tcaccgtggc tggcctcgcg caccgcgagg acggagcaga
```

```
21181 acatcgccgc cccgtcgagc ccgaacggga ccaggtactg ccagccgtcg ctcaggccga
21241 ggttctgctc gccgaagccg accagaccgt ggaaggagag cgccgcggcg accgcggcac
21301 agcagaagag caggacgtag gagaaggtgc cgtagagggc ctccttgcgc cggcggcgct
21361 cctccatgcg ctcccacgag tcgtccgcgc ccgcgtcctt ccccgaggaa cgcttgccgc
21421 gcgcgagcac cgccaccgcc gccagcatgc ccaggagcag cacggcgccc ggaagcagcc
21481 agttcagcga tatgtcggtc aatctcatct ggggtccctt gcagtgggat agggcgtaac
21541 gcccgccatg gtggcccaaa cccagcggcc ctcagggggt ttcggggcaa gaggccgcca
21601 agggagtgca gaggggtgc ccagggcggc gttctgctcg aactgccgcg tgaggggcgg
21661 gagttgagtt cgattaagac tacccgtacg ggtggttcca cggaaagttc ccgcgacgtg
21721 tgggggagtt gtggaacaac gcggacgatc ttagccgacg gcccgtcagc ccgccgcggc
21781 gagcagcttg tccacgcggt cggcgtcgca ggtccgcggg caggtggcgc aggtgtcctc
21841 ggggcgcagg gtgtagaaca tgcagcagct cgcccggtcc cgggtgtgca gcggggcgcc
21901 gtccggcccc ttcagttcac ggaaggccgc cttcccgacg tacggcctgg tggcgcccgg
21961 cagcagcagc tccagctcgt gccgcgcccg ctcctgctcg ccgagcaggt gggcgacgta
22021 ccacaggccc tcgacgacct cgtcggtcgc catcccccac agggcacgtc cgcgccgccg
22081 catccggggg ccgaatccgg ccagcaccgg ctccaggtgc tcggcgaccg cggcccgcac
22141 ctcggcgcgc agcgcctcct cgtcggcgac cacccgggcg ccgggcagcg cggcggcggg
22201 gtcgccgggg aggcaggcga aggacgcggc acgcaccgcc atccggccga gggggagccc
22261 ggcggccgtg cggtcgtagg acacgtgggc cgcggggtgg cggggcaccc ggcggtgcag
22321 gaaccacggc acggtgatca gcaggcaggc gggccacgcg taccggtgca ggccgaagct
22381 cgcgatcacg tccggccggg cccgctgccc gtagtcccgc agcacctggg cctcgtccca
22441 ggcgaggaag gccgccagct ccgtcccgcc cgcggcgagg gaggcggccg tgacccagcc
22501 gccgccgctc gggacctcgt cggccgcggc cagttcggtc acgcccagcc cggggaggac
22561 ctcggacagg cgcgtgtagg cgtccgtgag ggccgtacgg cgacaggca tgcgggacca
22621 ccgatccacg tcggaacaac gggctgctta aaggtaagcc ttaccttacc caatatgttc
22681 gggatttgaa ctgacggccc ctgcgcttat ggtgcttgac ggacaggtga caacccgccg
22741 taatgacccc aagtaccgac acgtcccaca ggaggacccg tgaagcacag cgcgcagggc
22801 gccgcaagaa cgggtgtggg tacgggtacg ggggcggcgg ccgaagcggt gcggatcccg
22861 gtgcagccgg gagcggcgga cagggcacgg ggccgggccc ccggcggtgc ggagccggac
22921 ggcccggcac gcggcgagca cacccacggc gaaccgccgg tgccccgccc gcgggccctg
22981 gtccagcgtt cctccgtgcg cggacagatc ctcgacgccc tgcgctccgc gctggtcacc
23041 ggtgaactgc ggcccggcga ggtctactcg gcaccggtgc tcggcgagcg gttcggcgtc
23101 tcggccaccc ccgtgcggga ggccatgcag cagctcgccc tggagggcgc cgtcgaggtc
23161 gtccccaacc ggggcttccg cgtcctcgaa cgcggtgacc gcgagctggc cgaactggcc
23221 gaggtccggg ccctgatcga ggtcccggtg tggctccggc tggcccgcac ggtgcccgcg
23281 gagcactggg cggagctgcg cccgctcgcc gagggcaccg tccgggccgc gtcctccggc
23341 tgcccggcca cctacgccga ggccgaccgc gccttccacc gcgccgcgct cgcgctggcc
23401 ggcaacgaac agctcgtccg gatcgccggt gacgtgcacc gccgcgccca gtggccgccg
23461 gccggctcgc cctccgtgcg cggccgggcc gacctggtcg ccgacgcgca ccagcacacc
23521 gcgctgctgg acgcgctgat cgccggcgac ttggacgtgg tgcgggccct ggtcggggac
23581 cacttcacag tcacgtcgtg agccggaggc ccggcagggc gcccttcacc cgggcgcccg
23641 gacggggccc ccggaccgga cgtgccccgg gtgggacgtg ccccgggccg gtcgtccgcc
23701 gggcttcgcc cggatcggcg ttcccgggcc ggggtcccgg gtcggcgttc ccgggccggg
23761 cgctcggacg gggcgggcca cggaccgcgc gccccgggcc cgggcgccct tgagcgaagg
23821 ggcccggacc ttcctgcccg cccctgctag cggacctcct gccgttccgt cggctcgtg
23881 ggcacctgct tccccggctt ctcggacttc cccgtggcgc gggtccggcc gcccctgcgg
23941 cgcccggggc gggaaccggt gtcgcggttg tgcagctcga tggccagggg agaggcggtg
24001 aagaccgacg agtacgtgcc gaccccacg ccgatcagca gcgcgagcgc gaagtcggtg
24061 agcgagtcgt cggccaggat cgccagcgag gccaggatga gcaccgcgcc catgccggtg
24121 ttgaccgtgc gcggcagggt ctgcaggatc gcgtcgttgg tgagccggtc gaacggtgcc
24181 ttccgctcct tgcccagcag ttcccggatg cggtcgaaga gcaccaccga gtcgttcacc
24241 gagtagccga tcacggtcag cagcgccgcc aggaacaccc cgtcgatcgg cttgcccagc
24301 caggcgaaca caccgaccag aatgaccacg tcgtgggcga gcgcgccgac cgccgccgtg
24361 ccgaacagca ggcggaaccg gaccgccagg tacgccagct gcgcggccag cgccaggccg
24421 agggcgatca gggcgtcccg gcgcagctcc tcgccgaggc tcgggccgat cagctcgtcg
24481 cggaccttct ccgtctcgcc gcccagttcg gcgacggcct tcgtgacggt cgccgcctcg
24541 gtgtcggtca gctcctccgt gcgcaccgtg aggtcgccgt cgccggagga ctggacgacg
24601 gcgcgcggga agcccgcgtc ggccagggcg tcccgcgccc ggtcgggatc gacctgggtg
24661 gcggtggagt actcgatgag ccgcccgccg gtgaactcga tgccgaagtt caggccccgc
24721 accaggatgc cggagcccgc caccacaagg acgatcagcg aggccgccag ccagcggcgc
24781 ggccggcgca tcaggaacgg gtcgcggcgc agcagggcgt cacggaccgg gccggtcgag
24841 gagatgccgg tgatacgggg gcggcggaag accgccgggc ggctcgcggc gaactcggcg
24901 agcacccggg tgatcaccag ggcgctgacc atcgaggcca gcacaccgat gccgagggtg
24961 acgccgaagc ccttgaccgg gccggaggcc aggaagaaca gcaggcggc ggcgatcagc
25021 gtggtgatgt tggagtcggc gatcgcgctg aaggcgctgc ggaaaccggc ggtcagcgcg
```

```
25081 gagcgcgtac tgggacgggt ccgggccgcc tgttcctcgc gggcccgctc gaagaccagg
25141 acgttggcgt ccaccgccat gccgatcgcc agcacgaagc ccgccaggcc cggcagggtc
25201 agggtcgcgc cgaccgcggc cagggcggcg taggagatca ggccgtagca gagcagggcc
25261 acggtcgcga gggcgcccat cagccggtag acgacgatga tgaacagcgc ggtcagcgcg
25321 gtgccgatga cggcggccca ggcgccggcg tcgatggcct cgtcgcccag ggtggcgccg
25381 atggtgcgct gctcgatggt ctcgaccggc accggcaggg cgccgccctt gatgagcagg
25441 gccagttcgc gcgcctcggc gtcgtcgaag gagccggtga tctgcgtgga gccgccggtg
25501 atgcccgccc cgcaggacac ggaggggtcg acctgcggcg aggagatgat cttgtcgtcg
25561 agcacgatgg cgacccggcg cgtggggtca ccggccgggt tgcaggccgc ctcgccggtc
25621 acctgcgccc agcggtcgct gccggagtcc ttgaagtcga cggtgacggt ccagccggcg
25681 ccgttctgct ggtcgaagcg ggcgtcggcg cccttgacgt cctgaccggt cagcgtggcg
25741 gccttcagcc gcagggactg cccggactcg tccggcagca cccgctcgcc gtcctcggcg
25801 gtctcggtcc tgccctcggt cgcgtcggcg ggcttctccg ccgcgccgag caccgagtgg
25861 acggtgagct gcgcggtgcg gccgagcacg tcggcggcct tcttcgggtc ctgtacgccc
25921 ggcagctcga cgacgacccg gttctcgccg gagcggacga tggtcggctc ggcgacgccg
25981 agggcgtcga tgcggccgcg cagcacctcc acggtgcggt cggtggcctc ccggtcggcc
26041 ttggtggtct cggtcgactt ggtctccagg acgatctggg tgccgccgcg caggtcgagc
26101 cccaggcgga ccggcatggt cagggcgatg gccacggccc cggccatcac ggccagcgcg
26161 acgagggcgc ggacgttcag ggaacgggac cgcgtacggg aacgggagcg ggaaggggga
26221 cgggagcgtt tcaagacggg cctccggcgg gcacgccgcg gcggcggccg agcgcggccg
26281 cggcggaaga agatgcaggg aaggagtggt gtctcaggtg ccggaggaga ccggtggcgc
26341 ccggccgctg tcgtggctcg cgcggccggg ggagaccggc accgggccgg cgcgaacga
26401 cgtccgcgcg tggccgcagg gggcggcgcc gcccccgagg tgggcggtgg tcgccgggtg
26461 ctggtccggt gccgtcggac gcccggccgg ctggtcgtgc cgggctcgag gctgggcggc
26521 gcagaccggc gtgcagccgt cgtcggcgtg gaggtcgagg gccgggtgcg gcacggccgt
26581 ggcggcggcc gccgcgaggg gaccgctcgg accgacgtgc gcaccggccg cgggccaggc
26641 cgggaccagg ggcagcagga cggtcagcag gacggtgagc acgctcgcga gccgtcgtcg
26701 catgcaccgt cccctttcct ggtggatgag cgggcctggc gggtggggga gcccggggct
26761 ccggccgggg agtgggcggc cagagccacg ggcggttcgg gggaccgtca gggctcgatg
26821 agcccggcgc ggatggcgta ccgggtcagc tccaggcggt cgcgcaggcc gagcttctgc
26881 agcaggttcg cccggtgccg ctggacggtc ttgatgctga tgaagaggat ctcggcgatc
26941 tccttcgagg agtggccctc ggcgacgagc ttgaggacct cctcctcgcg cggggtgagg
27001 atgtggtcgg aggtctcctc gccgtgccgc acccggtcga ggtagttgcg gatgagcgcg
27061 gtcaccgcgc ccgggtacag gaacggctcg tcgcgcatcg cggcgcggca ggcggcgacg
27121 aggtcccggt cggcgaccga cttgagcacg tatccgcagg cccccgactt cagcgcctgg
27181 aacaggtact gctcgttgtc gtgcatcgtc agcatcagga tgcgcagccc cggtttcagc
27241 gcggccagtt cgcgggccgc ctggaggccg gtgaggcgcg gcatcgcgat gtccaggacg
27301 gcgaggtcgg cctcgtgggc ccgcgccatg tcgatggcct ccgcgccgtc cccggcctcg
27361 gccaccacct ccaggtccgg ctcccggtcg aggatgaggc gcacaccgcg cgcgcaccagc
27421 gcgtggtcgt cggcgagcaa gatgcggatc ttcgaggggg cggacgccgg gggtgtgccg
27481 tacgcggacg tcgtggacgc ggacgtgggc gtcgtgggcg tcgtggacgc gccggcgtgg
27541 cggccgcgcg cggacggctc ggacggctcg ggcagggacg ggtcggacat ggtcggtgct
27601 acttcctggg aaggggtgcg gtgagccgga tgcgggtgcc ggcgccgggt gccgacgtga
27661 tgtcgagggc ggccccgatg agcagggccc gttcccgcat gccgcggatg ccggcgccct
27721 cgtgcgccgc ctcgatgccg cgcccgtcgt cggcgacggt gagcgtcacc gcggcgtcgg
27781 cgtgggccag gcccacctcc aggcgctcgg cgtccgcgtg gcgggcggcg ttggtcaggc
27841 tctcctgagc cacccggtag aggaccagtt cggtctcgtg gtccaggacc ggcagatcgg
27901 cgtcgaaacg gcgcaccacc cgcagaccgg tgtgggtggc gaagtcgtgc gtgagcgagg
27961 acagcgcgct gatcagaccg aggtcgtcca ggacaccggg ccgcagccgg cgcaccaggc
28021 ggcggacctc gtccaggctc tcccgggtga tctcctgcgc ctggtgcagc tcgtcgcgca
28081 gcggctcctc ggcgtcgtcc gcggcacgcc ccaggaccag caggatcgcg gtcatgctct
28141 gcccgacctc gtcgtgcagc tcctgggcga tgcggcgccg ttccgcctcc tgggccagca
28201 ggacgcgggc gctgctggtg gcccgctcgt tctccaggcg gtccagcatg gcgttgaagg
28261 tgcggatcag ctccgggacc tcgccgccgc cggagatcgg cagccgctgg ccggggcgca
28321 gcaggtcgac ggtggtcatc agcctggtga gccgctccag cggcgccagg ccccagcgca
28381 gcagggcacc gttggccacc agcatgacgc ccatgccgac cacgaggacg acggcctcgg
28441 tcagcaggat cggcacggag acggtcaccg gagcccacag cagcagcgcc gtggcgaagc
28501 ccagcaccac cgcgttgagc ccgaagatcc gccagaacag ggacaccggg ggtacgcctc
28561 cttcggcgac gtgttgattt ggtacgcctc tactgtcggt catcgcgcac cctcggtggg
28621 agctgcggcc caggtagcga ccgagcaagc gaccagagtg ctcctcgcgg accggtccgg
28681 tcgatggggc ccgacccca tttccggccc gcggcatacc catatggccg ccccggcgca
28741 ccgccgagtg gatcccgcac ccccgggcaa atgggtcctg cgcccgatgg gtttccccgg
28801 gcggcccggc cagggttgag gacgaccggg tccgatcggc ccggcactcg gcatacccgc
28861 agcaccttcc gagaaaagga ccactcatcg tgtcgctcga cgcctcccgc atcgaacccc
28921 gccccgagcg gctgaccgct cacgaggcgc gccagcgcct cgaacacgcc cgcaacaccc
```

```
28981 gcgtcaccca gcttcaggcc ctcgccgaga gcggccaggc ggacgaccag ctgatgtccg
29041 cgcagaaggc cgcgatcgag cgggtgctca aggagatcga cgaggccttc gcccgggtcg
29101 aggagggcac ctacggcgcc tgcctcggct gcggcaagcc cgtccccggg gagcgcctgg
29161 agatcctccc ctacacgcgg tactgcgtcg cctgccagcg ccgcgcggcc gcctgacgcc
29221 ccgtcccacc gtcccaccgt ccccgtccag taccgtccct cccgccctgc ccaaggggtg
29281 aagtggtgaa caaccagatc atcggcgacc gcgacacgcg tctgccgccc ctgtcgcccg
29341 aggacctcgc cgcgctccgg gacaacctcc gggagcaacg gctgttccgc gaggagcaac
```

//